

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2006, 19:25:44 ; Search time 328 Seconds
(without alignments)
93.356 Million cell updates/sec

Title: US-10-645-818-5
Perfect score: 42
Sequence: 1 gacgttggaagctgctgct.....accaactgagctaccaccgc 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4637633 seqs, 364532575 residues

Total number of hits satisfying chosen parameters: 9275266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	4524	7	US-11-075-185-61
2	31.4	74.8	75	7	US-11-088-257-1
3	24	57.1	171936	6	US-10-933-025-24
4	23.4	55.7	25	7	US-11-088-257-10
5	23.4	55.7	25	7	US-11-088-257-30
6	23	54.8	5071	6	US-10-750-185-31677
7	23	54.8	5071	6	US-10-750-185-31677
8	22.8	54.3	1155	6	US-10-750-185-51288
9	22.8	54.3	1155	6	US-10-750-623-51288
10	22.6	53.8	1717	6	US-10-750-185-45201
11	22.6	53.8	2652	6	US-10-750-623-45201
12	22.6	53.8	2652	6	US-10-750-185-30822
13	22.6	53.8	2652	6	US-10-750-623-30822
14	22.4	53.3	216	6	US-10-467-657-7779
15	22	52.4	2005	6	US-10-750-185-38017
16	22	52.4	2005	6	US-10-750-623-38017
17	21.8	51.9	598	6	US-10-750-185-2981
18	21.8	51.9	598	6	US-10-750-623-2981
19	21.8	51.9	1972	6	US-10-750-185-28416
20	21.8	51.9	1972	6	US-10-750-623-28416
21	21.8	51.9	2859	6	US-10-793-626-4402
22	21.8	51.9	3012	6	US-10-793-626-3707
23	21.8	51.9	3232	6	US-10-793-626-4217

ALIGNMENTS

RESULT 1

US-11-075-185-61
; Sequence 61, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 4524
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
; US-11-075-185-61

Query Match 100.0%; Score 42; DB 7; Length 4524;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTGCTGCTTACCAACTGAGCTACACCGC 42
|||||

Db 3830 GAGCTTGGGAAGCTGCTGCTTACCAACTGAGCTACACCGC 3871
|||||

RESULT 2

US-11-088-257-1/c
; Sequence 1, Application US/11088257
; Publication No. US20050260636A1
; GENERAL INFORMATION:
; APPLICANT: CELLEX INC
; TITLE OF INVENTION: Methods for Detection of Biological Organisms Using tRNA
; FILE REFERENCE: CELLEX--0006
; CURRENT APPLICATION NUMBER: US/11/088,257
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: US 60/555,683
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 30

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 75
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Primer for bacteria
US-11-088-257-1

Query Match          74.8%; Score 31.4; DB 7; Length 75;
Best Local Similarity 85.4%; Pred. No. 0.00091;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCAACG 41
Db 41 GACCTTGGCAAGTGTGCTCTTACCAACTGAGCTATTCCG 1

RESULT 3
US-10-933-025-24
; Sequence 24, Application US/10933025
; Publication No. US20050265987A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN, STEVEN
; APPLICANT: HEMMERICH, STEFAN
; APPLICANT: TOMITA, MEGUMI
; TITLE OF INVENTION: Sulfotransferases and methods of use
; FILE REFERENCE: UCAL-230CON
; CURRENT APPLICATION NUMBER: US/10/933,025
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 171936
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-933-025-24

Query Match          57.1%; Score 24; DB 6; Length 171936;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 GCTTGGGAAGCTCGTCTTACCAACTGAGCTACCAACG 42
Db 37347 GCTGGCGTGGCTGGCGCTTTACCACTGAGCCACCGC 37386

RESULT 4
US-11-088-257-10
; Sequence 10, Application US/11088257
; Publication No. US20050260636A1
; GENERAL INFORMATION:
; APPLICANT: CELLEX INC
; FILE REFERENCE: CELLEX--0006
; CURRENT APPLICATION NUMBER: US/11/088,257
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: US 60/555,683
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 25
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Primer for bacteria
US-11-088-257-10

Query Match          55.7%; Score 23.4; DB 7; Length 25;
Best Local Similarity 96.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 AAGCTCGTGTCTTACCAACTGAGCT 34
Db 1 AAGCTCGTGTCTTACCAACTGAGCT 25

RESULT 5
US-11-088-257-30
; Sequence 30, Application US/11088257
; Publication No. US20050260636A1
; GENERAL INFORMATION:
; APPLICANT: CELLEX INC
; FILE REFERENCE: CELLEX--0006
; CURRENT APPLICATION NUMBER: US/11/088,257
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: US 60/555,683
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 25
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Primer for bacteria
US-11-088-257-30

Query Match          55.7%; Score 23.4; DB 7; Length 25;
Best Local Similarity 96.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 AAGCTCGTGTCTTACCAACTGAGCT 34
Db 1 AAGCTCGTGTCTTACCAACTGAGCT 25

RESULT 6
US-10-750-185-31677/c
; Sequence 31677, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31677
; LENGTH: 5071
; TYPE: DNA
; ORGANISM: Bovine 19866880454815
US-10-750-185-31677

Query Match          54.8%; Score 23; DB 6; Length 5071;
Best Local Similarity 74.4%; Pred. No. 4.7;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

QY 3 GCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCG 41
|||||
Db 1391 GCATTGCAAGCGGATCTTTACCAACTGAGCTATCACAG 1353

RESULT 7

US-10-750-623-31677/c
; Sequence 31677, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31677
; LENGTH: 5071
; TYPE: DNA
; ORGANISM: Bovine 19866880733794
US-10-750-623-31677

Query Match 54.8%; Score 23; DB 6; Length 5071;
Best Local Similarity 74.4%; Pred. No. 4.7;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 GCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCG 41
|||||
Db 1391 GCATTGCAAGCGGATCTTTACCAACTGAGCTATCACAG 1353

RESULT 8

US-10-750-185-51288
; Sequence 51288, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 51288
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Bovine 19866880733794
US-10-750-185-51288

Query Match 54.3%; Score 22.8; DB 6; Length 1155;
Best Local Similarity 71.4%; Pred. No. 4.5;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCG 42
|||||
Db 1037 GAGCATCTCAAGCTCGTCTTACCAAGAGAACCACTGC 1078

RESULT 9

US-10-750-623-51288
; Sequence 51288, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 51288
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Bovine 19866880733794
US-10-750-623-51288

Query Match 54.3%; Score 22.8; DB 6; Length 1155;
Best Local Similarity 71.4%; Pred. No. 4.5;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCG 42
|||||
Db 1037 GAGCATCTCAAGCTCGTCTTACCAAGAGAACCACTGC 1078

RESULT 10

US-10-750-185-45201
; Sequence 45201, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 45201
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Bovine 19866881206092
US-10-750-185-45201

Query Match 53.8%; Score 22.6; DB 6; Length 1717;
Best Local Similarity 75.7%; Pred. No. 5.7;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 TTGGGAAGCTCGTCTTACCAACTGAGCTACACCG 41
|||||
Db 1601 TTGGCAGGCGGTTCTTTACCACTGAGCCACACAG 1637

RESULT 11


```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38017
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Bovine 19866880521244
US-10-750-185-38017

```

```

Query Match      52.4%; Score 22; DB 6; Length 2005;
Best Local Similarity 73.7%; Pred.No. 10;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy      3  GCTTGGGAAGCTCGTCTTACCACTGAGCTACACC 40
          |||||
Db      579 GCATTGCAGGCTGATCTTTACCACTGAGCTATCAGC 616

```

Search completed: January 10, 2006, 21:48:32
Job time : 329 secs

This page blank (uspto)

Result No.	Score	Query		ID	Description
		Match	Length		
1	42	100.0	4575	3	US-09-903-540-657
2	32.4	77.1	258	3	US-09-540-236-104
3	32.4	77.1	549	3	US-09-252-991A-14851
4	32.4	77.1	690	3	US-09-252-991A-15110
5	32.4	77.1	5559	2	US-08-287-442-3
6	32.4	77.1	5559	2	US-08-459-701-3
7	32.4	77.1	5559	2	US-08-460-298-3
8	32.4	77.1	5559	2	US-08-459-174-3
9	32.4	77.1	5559	6	US-08-459-174-3
10	32.4	77.1	10763	2	US-08-761-258-1
11	32.4	77.1	10763	2	US-08-977-306-1
12	32.4	77.1	31147	3	US-09-596-002-25
13	32.4	77.1	63563	3	US-09-596-002-33
14	32.4	77.1	1830121	3	US-09-557-884-1
15	32.4	77.1	1830121	3	US-09-557-884-1
16	32.4	77.1	1830121	3	US-09-557-884-1
17	29.8	71.0	339	3	US-09-543-681A-3919
18	29.8	71.0	5595	3	US-09-221-017B-650
19	29.8	71.0	18431	3	US-09-221-017B-1090
20	29.8	71.0	640681	3	US-09-790-988-1
21	28.2	67.1	282	3	US-09-328-352-726
22	28.2	67.1	7035	3	US-09-903-540-878
23	27.2	64.8	603	2	US-08-413-614-85
24	27.2	64.8	603	2	US-08-413-614-85


```
; NAME/KEY: misc feature
; LOCATION: 4616..4691
; OTHER INFORMATION: /note= "glyw, transcribed right to
; OTHER INFORMATION: left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4731..5318
; OTHER INFORMATION: /note= "ORF 3, transcribed right to
; OTHER INFORMATION: left"
; US-08-287-442-3

Query Match      77.1%; Score 32.4; DB 2; Length 5559;
Best Local Similarity 85.7%; Pred. No. 0.0037;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGCTTACCAACTGAGCTACCAACGCGC 42
Db 4650 GACCTTGGCAAGGTCGTCTTACCAACTGAGCTATTCCGCGC 4691

RESULT 6
US-08-459-701-3
; Sequence 3, Application US/08459701
; Patent No. 5686282
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas D.
; APPLICANT: Lam, Stephen T.
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Stein, Jeffrey I.
; APPLICANT: Howell, Charles R.
; APPLICANT: Becker, J. Ole
; NUMBER OF SEQUENCES: 9
; TITLE OF INVENTION: Gene Activating Element
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,701
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/287,442
; FILING DATE: 08-AUG-1994
; APPLICATION NUMBER: US 08/087,636
; FILING DATE: 01-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,284
; FILING DATE: 02-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/570,184
; FILING DATE: 08-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; TELEPHONE: 919-541-8614
; TELECOMMUNICATION INFORMATION:
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5559 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
; STRAIN: CGA67356
; INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
; INDIVIDUAL ISOLATE: fragment
; IMMEDIATE SOURCE:
; CLONE: PCIB137
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 210..1688
; OTHER INFORMATION: /note= "ORF 1, transcribed left to
; OTHER INFORMATION: right"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1906..3633
; OTHER INFORMATION: /note= "ORF 2, transcribed left to
; OTHER INFORMATION: right"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4616..4691
; OTHER INFORMATION: /note= "glyw, transcribed right to
; OTHER INFORMATION: left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4731..5318
; OTHER INFORMATION: /note= "ORF 3, transcribed right to
; OTHER INFORMATION: left"
; US-08-459-701-3

Query Match      77.1%; Score 32.4; DB 2; Length 5559;
Best Local Similarity 85.7%; Pred. No. 0.0037;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGCTTACCAACTGAGCTACCAACGCGC 42
Db 4650 GACCTTGGCAAGGTCGTCTTACCAACTGAGCTATTCCGCGC 4691

RESULT 7
US-08-460-298-3
; Sequence 3, Application US/08460298
; Patent No. 5686283
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas D.
; APPLICANT: Lam, Stephen T.
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Stein, Jeffrey I.
; APPLICANT: Howell, Charles R.
; APPLICANT: Becker, J. Ole
; TITLE OF INVENTION: Gene Activating Element
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,298
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/460,298
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```


OTHER INFORMATION: /note= "ORF 3, transcribed right to
OTHER INFORMATION: left"
US-08-459-174-3

Query Match 77.1%; Score 32.4; DB 2; Length 5559;
Best Local Similarity 85.7%; Pred. No. 0.0037;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCAACGC 42
Db 4650 GACCTTGGCAAGCTCGTCTTACCAACTGAGCTATTCCCGC 4691

RESULT 9

PCT-US93-06300A-3

Sequence 3, Application PC/TUS9306300A

GENERAL INFORMATION:

APPLICANT: CIBA-GEIGY AG

APPLICANT: Klybeckstrasse 141

APPLICANT: 4002 Basle

APPLICANT: Switzerland

APPLICANT: 125 Tradescant Road

APPLICANT: Chapel Hill, NC 27514

APPLICANT: USA

APPLICANT: 8900 Jeanew Court

APPLICANT: Raleigh, NC 27613

APPLICANT: USA

APPLICANT: Hill, Dwight Steven

APPLICANT: 311 Melanie Lane

APPLICANT: Cary, NC 27511

APPLICANT: USA

APPLICANT: Stein, Jeffrey I.

APPLICANT: 3725 Surry Trail

APPLICANT: Hillsborough, NC 27278

APPLICANT: USA

APPLICANT: Howell, Charles R.

APPLICANT: 805 Avondale

APPLICANT: Bryan, TX 77802

APPLICANT: USA

APPLICANT: Becker, J. Ole

APPLICANT: 6164 Osevego

APPLICANT: Riverside, CA 92506

APPLICANT: USA

APPLICANT: Ligon, James M.

APPLICANT: 120 Marquette Drive

APPLICANT: Cary, NC 27513

APPLICANT: USA

TITLE OF INVENTION: Gene Activating Element

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: New York

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06300A

FILING DATE: 02-JUL-1993

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/908,284

FILING DATE: 02-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: S-18210/A/CGC1506/PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5559 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudomonas fluorescens
STRAIN: CGA267356
INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
INDIVIDUAL ISOLATE: fragment
IMMEDIATE SOURCE:
CLONE: pCIB137
FEATURE:
NAME/KEY: misc feature
LOCATION: 210..1688
OTHER INFORMATION: /note= "ORF 1, transcribed left to
OTHER INFORMATION: right"
FEATURE:
NAME/KEY: misc feature
LOCATION: 1906..3633
OTHER INFORMATION: /note= "ORF 2, transcribed left to
OTHER INFORMATION: right"
FEATURE:
NAME/KEY: misc feature
LOCATION: 4616..4691
OTHER INFORMATION: /note= "glyW, transcribed right to
OTHER INFORMATION: left"
FEATURE:
NAME/KEY: misc feature
LOCATION: 4731..5318
OTHER INFORMATION: /note= "ORF 3, transcribed right to
OTHER INFORMATION: left"
PCT-US93-06300A-3

Query Match 77.1%; Score 32.4; DB 6; Length 5559;
Best Local Similarity 85.7%; Pred. No. 0.0037;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCAACGC 42
Db 4650 GACCTTGGCAAGCTCGTCTTACCAACTGAGCTATTCCCGC 4691

RESULT 10

US-08-761-258-1

Sequence 1, Application US/08761258

Patent No. 5756087

GENERAL INFORMATION:

APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight S.

APPLICANT: Lam, Stephen T.

APPLICANT: Gaffney, Thomas D.

APPLICANT: Torkewitz, Nancy

TITLE OF INVENTION: Genetically Modified Pseudomonas Strains

WITH ENHANCED BIOCONTROL ACTIVITY

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 520 White Plains Road, P.O. Box 2005

CITY: Tarrytown

STATE: NY

COUNTRY: USA

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/761,258
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meigs, J. Timothy
/ REGISTRATION NUMBER: 38,241
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (919) 541-8587
/ TELEFAX: (919) 541-8689
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10763 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Pseudomonas fluorescens
/ STRAIN: CGA267356 (aka MCG134 and aka BL915)
/ IMMEDIATE SOURCE:
/ CLONE: Plasmid pE11
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 210..1688
/ OTHER INFORMATION: /product= "methyltransferase"
/ OTHER INFORMATION: /note= "Coding sequence for methyltransferase has homology to
/ OTHER INFORMATION: the cheR and frzF genes from E. coli and Myxococcus xanthus,
/ OTHER INFORMATION: respectively."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1906..3633
/ OTHER INFORMATION: /product= "sensor kinase"
/ OTHER INFORMATION: /note= "Coding sequence for sensor kinase has homology to the
/ OTHER INFORMATION: rcsC, frzE, and bvgS genes of E. coli, M. Xanthus, and
/ OTHER INFORMATION: Borditella pertussis, respectively."
/ FEATURE:
/ NAME/KEY: misc_RNA
/ LOCATION: complement (4616..4691)
/ OTHER INFORMATION: /product= "tRNA"
/ OTHER INFORMATION: /note= "(complementary DNA strand) Homology to glyW from E.
/ OTHER INFORMATION: coli."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (4731..5318)
/ OTHER INFORMATION: /product=
/ OTHER INFORMATION: "CDP-diacylglycerol-glycerol-3-phosphate-3-phosph
/ OTHER INFORMATION: atidyltrans."
/ OTHER INFORMATION: /note= "Coding sequence for
/ OTHER INFORMATION: CDP-diacylglycerol-glycerol-3-phosphate-3-phosphatidyltransfe
/ OTHER INFORMATION: se has homology to pgsA."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (5574..7397)
/ OTHER INFORMATION: /product= "UVR exonuclease subunit
/ OTHER INFORMATION: C"
/ OTHER INFORMATION: /note= "Coding sequence for UVR exonuclease subunit C has
/ OTHER INFORMATION: homology to uvrC."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (7400..8041)
/ OTHER INFORMATION: /function= "response
/ OTHER INFORMATION: regulator/transcription activator"
/ OTHER INFORMATION: /product= "gacA (aka gafa)"
/ OTHER INFORMATION: /note= "Coding sequence for gacA (aka gafa) has homology to b
/ OTHER INFORMATION: uvrY and gacA genes of E. coli and Ps. fluorescens,
/ OTHER INFORMATION: respectively."
/ OTHER INFORMATION:
/ US-08-761-258-1
/ Query Match 77.1%; Score 32.4; DB 2; Length 10763;
/
```

```
Best Local Similarity 85.7%; Pred. No. 0.0043; 6; Indels 0; Gaps 0;
Matches 36; Conservative 0; Mismatches 0;
QY 1 GAGCTTGGGAAGCTGCTGCTCTACCACTGAGTACCACGC 42
DB 4650 GACCTTGGCAAGTGTGCTCTACCACTGAGTATTCCCGC 4691

RESULT 11
US-08-977-306-1
; Sequence 1, Application US/08977306
; Patent No. 5955348
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Gaffney, Thomas D.
; APPLICANT: Torkewitz, Nancy
; APPLICANT: Stafford, Jill M.
; TITLE OF INVENTION: Genetically Modified Pseudomonas Strains
; TITLE OF INVENTION: with Enhanced Biocontrol Activity
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5955348artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,306
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10763 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
; STRAIN: CGA267356 (aka MCG134 and aka BL915)
; IMMEDIATE SOURCE:
; CLONE: Plasmid pE11
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 210..1688
; OTHER INFORMATION: /product= "methyltransferase"
; OTHER INFORMATION: /note= "Coding sequence for methyltransferase has homology to
; OTHER INFORMATION: the cheR and frzF genes from E. coli and Myxococcus xanthus,
; OTHER INFORMATION: respectively."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1906..3633
; OTHER INFORMATION: /product= "sensor kinase"
; OTHER INFORMATION: /note= "Coding sequence for sensor kinase has homology to the
; OTHER INFORMATION: rcsC, frzE, and bvgS genes of E. coli, M. Xanthus, and
; OTHER INFORMATION: Borditella pertussis, respectively."
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: complement (4616..4691)
; OTHER INFORMATION: /product= "tRNA"
; OTHER INFORMATION: /note= "(complementary DNA strand) Homology to glyW from E.
; OTHER INFORMATION: coli."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (4731..5318)
; OTHER INFORMATION: /product=
; OTHER INFORMATION: "CDP-diacylglycerol-glycerol-3-phosphate-3-phosph
; OTHER INFORMATION: atidyltrans."
; OTHER INFORMATION: /note= "Coding sequence for
; OTHER INFORMATION: CDP-diacylglycerol-glycerol-3-phosphate-3-phosphatidyltransfe
; OTHER INFORMATION: se has homology to pgsA."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (5574..7397)
; OTHER INFORMATION: /product= "UVR exonuclease subunit
; OTHER INFORMATION: C"
; OTHER INFORMATION: /note= "Coding sequence for UVR exonuclease subunit C has
; OTHER INFORMATION: homology to uvrC."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (7400..8041)
; OTHER INFORMATION: /function= "response
; OTHER INFORMATION: regulator/transcription activator"
; OTHER INFORMATION: /product= "gacA (aka gafa)"
; OTHER INFORMATION: /note= "Coding sequence for gacA (aka gafa) has homology to b
; OTHER INFORMATION: uvrY and gacA genes of E. coli and Ps. fluorescens,
; OTHER INFORMATION: respectively."
; OTHER INFORMATION:
; US-08-761-258-1
; Query Match 77.1%; Score 32.4; DB 2; Length 10763;
;
```

```

RESULT 13
US-09-596-002-33/c
; Sequence 33, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 63563
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte template ID No. 6632636 33
; PUBLICATION INFORMATION:
US-09-596-002-33

Query Match 77.1%; Score 32.4; DB 3; Length 63563;
Best Local Similarity 85.7%; Pred.No. 0.0084; 6; Indels 0; Gaps
Matches 36; Conservative 0; Mismatches 6;

QY 1 GAGCTTGGGAAGTCGTGCTCTACCAACTGAGCTACCAACCG 42
    |||||
DB 1051 GACCTTGGCAAGTCGTGCTCTACCAACTGAGCTATTCCGCG 1010
    |||||

RESULT 14
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: the Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439

```

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 77.1%; Score 32.4; DB 3; Length 1830121;
Best Local Similarity 85.7%; Pred. No. 0.014;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTGCTGCTCTACCAACTGAGCTACCAACGC 42
DB 90995 GACCTTGGCAAGGTCGTGCTCTACCAACTGAGCTATTCCGC 90954

RESULT 15
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: FB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-643-990A-1

Query Match 77.1%; Score 32.4; DB 3; Length 1830121;
Best Local Similarity 85.7%; Pred. No. 0.014;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTGCTGCTCTACCAACTGAGCTACCAACGC 42
DB 90995 GACCTTGGCAAGGTCGTGCTCTACCAACTGAGCTATTCCGC 90954
Search completed: January 10, 2006, 19:28:15
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2006, 18:22:31 ; Search time 815 Seconds
(without alignments)
426.152 Million cell updates/sec

Title: US-10-645-818-5

Perfect score: 42

Sequence: 1 gagcttgggaagctcgtgct.....accaactgagctaccaccgc 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.Main.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	100.0	42	8	US-10-645-818-5
2	42	100.0	74	8	US-10-645-818-12
3	42	100.0	257	8	US-10-645-818-7
4	42	100.0	360	8	US-10-645-818-3
5	42	100.0	360	8	US-10-645-818-4
6	42	100.0	1647	8	US-10-645-818-1
7	32.4	77.1	1657	3	US-09-815-242-3781
8	32.4	77.1	1960	9	US-10-795-159-201
9	32.4	77.1	5436	9	US-10-795-159-492
10	32.4	77.1	28871	9	US-10-795-159-575
11	32.4	77.1	31147	7	US-10-672-787-25
12	32.4	77.1	63563	7	US-10-672-787-33
13	32.4	77.1	417576	9	US-10-795-159-684
14	32.4	77.1	908766	9	US-10-795-159-685
15	32.4	77.1	1830121	7	US-10-329-670-1
16	32.4	77.1	1830121	8	US-10-158-865-1
17	32.4	77.1	1830121	9	US-10-981-687-1
18	29.8	71.0	5595	5	US-10-194-163-650
19	29.8	71.0	18431	5	US-10-194-163-1090
20	29.8	71.0	640681	3	US-09-790-988-1
21	29.8	71.0	9025608	6	US-10-156-761-1
22	29.8	71.0	9025608	6	US-10-156-761-1
23	28.2	67.1	4927	9	US-10-915-740A-14

24	28.2	67.1	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
25	28.2	67.1	2731748	7	US-10-297-465A-1	Sequence 1, Appli
C 26	27.2	64.8	603	3	US-09-863-086-85	Sequence 85, Appl
C 27	27.2	64.8	603	3	US-09-863-086-86	Sequence 85, Appl
C 28	27.2	64.8	603	7	US-10-672-238-85	Sequence 85, Appl
C 29	27.2	64.8	603	7	US-10-672-238-86	Sequence 85, Appl
C 30	27.2	64.8	654	3	US-09-863-086-89	Sequence 89, Appl
C 31	27.2	64.8	654	7	US-10-672-238-89	Sequence 89, Appl
C 32	27.2	64.8	20844	9	US-10-915-740A-8	Sequence 8, Appli
C 33	27.2	64.8	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap
C 34	27.2	64.8	3309400	3	US-09-738-626-1	Sequence 1, Appli
C 35	26.6	63.3	3309400	3	US-09-738-626-1	Sequence 1, Appli
C 36	26	61.3	2731748	7	US-10-297-465A-1	Sequence 1, Appli
37	25.8	61.4	104245	6	US-10-160-807-4	Sequence 4, Appli
38	25.8	61.4	104245	7	US-10-655-847-4	Sequence 4, Appli
C 39	25.8	61.4	170245	7	US-10-717-597-322	Sequence 322, App
C 40	25.6	61.0	808	3	US-09-931-486-131	Sequence 131, App
C 41	25.6	61.0	808	3	US-09-931-486-132	Sequence 131, App
C 42	25.6	61.0	808	9	US-10-895-114-131	Sequence 131, App
C 43	25.6	61.0	808	9	US-10-895-114-132	Sequence 132, App
C 44	25.6	61.0	809	3	US-09-931-486-154	Sequence 154, App
C 45	25.6	61.0	809	9	US-10-895-114-154	Sequence 154, App

ALIGNMENTS

RESULT 1

US-10-645-818-5

; Sequence 5, Application US/10645818

; Publication No. US20040247620A1

; GENERAL INFORMATION:

; APPLICANT: Kosan Biosciences, Inc.

; APPLICANT: Julien, Bryan

; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE

; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS

; TITLE OF INVENTION: BACTERIOPHAGE MX9

; FILE REFERENCE: 300622009940

; CURRENT APPLICATION NUMBER: US/10/645,818

; CURRENT FILING DATE: 2003-08-20

; PRIOR APPLICATION NUMBER: US 60/405,196

; PRIOR FILING DATE: 2002-08-21

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 42

; TYPE: DNA

; ORGANISM: Bacteriophage MX9

US-10-645-818-5

Query Match 100.0%; Score 42; DB 8; Length 42;

Best Local Similarity 100.0%; Pred. No. 3.4e-07;

Mismatches 0; Indels 0; Gaps 0;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42

Db 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42

RESULT 2

US-10-645-818-12/c

; Sequence 12, Application US/10645818

; Publication No. US20040247620A1

; GENERAL INFORMATION:

; APPLICANT: Kosan Biosciences, Inc.

; APPLICANT: Julien, Bryan

; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE

; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS

; TITLE OF INVENTION: BACTERIOPHAGE MX9

; FILE REFERENCE: 300622009940

; CURRENT APPLICATION NUMBER: US/10/645,818

; CURRENT FILING DATE: 2003-08-20

; PRIOR APPLICATION NUMBER: US 60/405,196

```
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 74
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-10-645-818-12

Query Match      100.0%; Score 42; DB 8; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 42
    |||||
Db 42 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 1

RESULT 3
US-10-645-818-7
; Sequence 7, Application US/10645818
; Publication No. US20040247620A1
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE
; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS
; TITLE OF INVENTION: BACTERIOPHAGE MX9
; FILE REFERENCE: 300622009940
; CURRENT APPLICATION NUMBER: US/10/645,818
; PRIOR FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/405,196
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Bacteriophage MX9
US-10-645-818-7

Query Match      100.0%; Score 42; DB 8; Length 257;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 42
    |||||
Db 115 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 156

RESULT 4
US-10-645-818-3
; Sequence 3, Application US/10645818
; Publication No. US20040247620A1
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE
; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS
; TITLE OF INVENTION: BACTERIOPHAGE MX9
; FILE REFERENCE: 300622009940
; CURRENT APPLICATION NUMBER: US/10/645,818
; PRIOR FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/405,196
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Bacteriophage MX9
US-10-645-818-3

Query Match      100.0%; Score 42; DB 8; Length 360;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 42
    |||||
Db 163 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 204

RESULT 5
US-10-645-818-4
; Sequence 4, Application US/10645818
; Publication No. US20040247620A1
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE
; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS
; TITLE OF INVENTION: BACTERIOPHAGE MX9
; FILE REFERENCE: 300622009940
; CURRENT APPLICATION NUMBER: US/10/645,818
; PRIOR FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/405,196
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Bacteriophage MX9
US-10-645-818-4

Query Match      100.0%; Score 42; DB 8; Length 360;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 42
    |||||
Db 165 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 206

RESULT 6
US-10-645-818-1
; Sequence 1, Application US/10645818
; Publication No. US20040247620A1
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; TITLE OF INVENTION: TRANSFORMATION SYSTEM BASED ON THE
; TITLE OF INVENTION: INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS
; TITLE OF INVENTION: BACTERIOPHAGE MX9
; FILE REFERENCE: 300622009940
; CURRENT APPLICATION NUMBER: US/10/645,818
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US 60/405,196
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1647
; TYPE: DNA
; ORGANISM: Bacteriophage MX9
US-10-645-818-1

Query Match      100.0%; Score 42; DB 8; Length 1647;
Best Local Similarity 100.0%; Pred. No. 5.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 42
    |||||
Db 1394 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACGC 1435
```



```
RESULT 7
US-09-815-242-3781/c
; Sequence 3781, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3781
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(657)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-242-3781

Query Match 77.1%; Score 32.4; DB 3; Length 657;
Best Local Similarity 85.7%; Pred. No. 0.004;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
Db 579 GACCTTGGCAAGGTCGTCTTACCAACTGAGCTATTCCCGC 538

RESULT 8
US-10-795-159-201
; Sequence 201, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 201
; LENGTH: 1960
; TYPE: DNA
; ORGANISM: H. influenzae
US-10-795-159-201

Query Match 77.1%; Score 32.4; DB 3; Length 657;
Best Local Similarity 85.7%; Pred. No. 0.004;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
Db 579 GACCTTGGCAAGGTCGTCTTACCAACTGAGCTATTCCCGC 538

RESULT 9
US-10-795-159-492/c
; Sequence 492, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 492
; LENGTH: 5436
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (101)..(1101)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5156)..(5156)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5265)..(5265)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5294)..(5294)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5391)..(5392)
; OTHER INFORMATION: n = a, c, g, or t
US-10-795-159-492

Query Match 77.1%; Score 32.4; DB 9; Length 5436;
Best Local Similarity 85.7%; Pred. No. 0.0052;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
Db 5153 GACCTTGGCAAGGTCGTCTTACCAACTGAGCTATTCCCGC 5112

RESULT 10
US-10-795-159-575/c
; Sequence 575, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 575
; LENGTH: 1960
; TYPE: DNA
; ORGANISM: H. influenzae
US-10-795-159-575
```

```
; LENGTH: 28871
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (82)..(82)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5660)..(5660)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28833)..(28833)
; OTHER INFORMATION: n = a, c, g, or t
;
US-10-795-159-575

Query Match      77.1%; Score 32.4; DB 9; Length 28871;
Best Local Similarity 85.7%; Pred. No. 0.0065;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42
   |||||
Db 9246 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGC 9205

RESULT 11
US-10-672-787-25/c
; Sequence 25, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025C1
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 2003-09-26
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 31147
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
;
US-10-672-787-25

Query Match      77.1%; Score 32.4; DB 7; Length 31147;
Best Local Similarity 85.7%; Pred. No. 0.0065;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42
   |||||
Db 13927 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGC 13886

RESULT 12
US-10-672-787-33/c
; Sequence 33, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025C1
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
;
US-10-672-787-33

; SEQ ID NO 33
; LENGTH: 63563
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
;
US-10-672-787-33

Query Match      77.1%; Score 32.4; DB 7; Length 63563;
Best Local Similarity 85.7%; Pred. No. 0.0072;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42
   |||||
Db 1051 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGC 1010

RESULT 13
US-10-795-159-684/c
; Sequence 684, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 684
; LENGTH: 417576
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (93603)..(93603)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (93620)..(93620)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (93622)..(93625)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (138581)..(138581)
; OTHER INFORMATION: n is a, c, g, or t
;
US-10-795-159-684

Query Match      77.1%; Score 32.4; DB 9; Length 417576;
Best Local Similarity 85.7%; Pred. No. 0.0091;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACCAACGC 42
   |||||
Db 76404 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGC 76363

RESULT 14
US-10-795-159-685/c
; Sequence 685, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PERL Program
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685
; LENGTH: 908766
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9223)..(9223)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (39640)..(39640)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292404)..(292404)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292495)..(292495)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908604)..(908641)
; OTHER INFORMATION: n is a, c, g, or t
; US-10-795-159-685
;
; Query Match 77.1%; Score 32.4; DB 9; Length 908766;
; Best Local Similarity 85.7%; Pred. No. 0.01;
; Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
;
; Qy 1 GACCTGGGAAGTCGTCCTACCAACTGAGCTACACCGC 42
; |||||
; Db 748266 GACCTGGGAAGTCGTCCTACCAACTGAGCTATTCGCGC 748225
;
; RESULT 15
; US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/436,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
```

```
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match 77.1%; Score 32.4; DB 7; Length 1830121;
Best Local Similarity 85.7%; Pred. No. 0.011;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GAGCTTGGGAAGCTGCTGCTTACCAACTGAGCTACCGC 42
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90995 GACCTTGGCAAGTGTGCTCTACCAACTGAGCTATTCGCGC 90954
```

Search completed: January 10, 2006, 21:43:12
Job time : 836 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2006, 14:16:17 ; Search time 2184 Seconds
(without alignments)
1093.143 Million cell.updates/sec

Title: US-10-645-818-5
Perfect score: 42
Sequence: 1 gagcttgggaagctgtgct.....accactgagctaccaccgc 42

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.in.*
- 3: gb.env.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pr.*
- 9: gb.ro.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.htg.*
- 15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	360	1	AY297771 Myxococ
2	42	100.0	4575	6	AR619262 Sequence
3	42	100.0	4581	7	AY247757 Bacteriop
4	36.2	86.2	300162	1	AE016929 Bacteroid
5	36.2	86.2	301933	1	AE017174 Porphyrom
6	34.6	82.4	9890	1	AE001788 Thermotog
7	32.4	77.1	76	1	ECOGLYW
8	32.4	77.1	76	1	ECOTRG3
9	32.4	77.1	258	6	AR449440 Sequence
10	32.4	77.1	423	1	ECTRNAG3
11	32.4	77.1	480	11	AY657813 Synthetic
12	32.4	77.1	490	1	AF274339 Pseudomon
13	32.4	77.1	500	1	ECTRNGCL
14	32.4	77.1	636	1	HIHPIATT
15	32.4	77.1	890	1	ECGLYWG
16	32.4	77.1	956	1	ECOPGSA
17	32.4	77.1	1960	6	CQ872588 Sequence
18	32.4	77.1	2591	1	PPAU4950 Pseudomon

C 19	32.4	77.1	3286	1	AY379549
C 20	32.4	77.1	3291	1	AY379548
C 21	32.4	77.1	5436	6	CQ872879
C 22	32.4	77.1	5559	6	I66429
C 23	32.4	77.1	5559	6	I73354
C 24	32.4	77.1	5559	6	I73361
C 25	32.4	77.1	5559	6	I81321
C 26	32.4	77.1	8232	1	AB088753
C 27	32.4	77.1	8493	1	AE013664
C 28	32.4	77.1	9117	1	PSEMETH
C 29	32.4	77.1	10029	1	AE013847
C 30	32.4	77.1	10668	1	U32844
C 31	32.4	77.1	10686	1	AE004687
C 32	32.4	77.1	10763	6	AR008973
C 33	32.4	77.1	10763	6	AR075314
C 34	32.4	77.1	11350	1	U32694
C 35	32.4	77.1	12542	1	AE004708
C 36	32.4	77.1	13409	1	U32698
C 37	32.4	77.1	19682	1	D90832
C 38	32.4	77.1	20418	1	AE008904
C 39	32.4	77.1	22099	1	AE008786
C 40	32.4	77.1	28871	6	CQ872962
C 41	32.4	77.1	31147	6	AR408746
C 42	32.4	77.1	31147	6	AX067450
C 43	32.4	77.1	63563	6	AR408754
C 44	32.4	77.1	63563	6	AX067458
C 45	32.4	77.1	105027	1	AJ617740

ALIGNMENTS

RESULT 1
AY297771
LOCUS
DEFINITION Myxococcus xanthus strain DZ1 linear BCT 27-OCT-2003
attB2 site, complete sequence.
ACCESSION AY297771 GI:34329321
VERSION
KEYWORDS
SOURCE Myxococcus xanthus
ORGANISM Myxococcus xanthus
Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
Cystobacterineae; Myxococcaceae; Myxococcus.
REFERENCE 1 (bases 1 to 360)
AUTHORS Julien,B.
TITLE Characterization of the integrase gene and attachment site for the
Myxococcus xanthus bacteriophage Mx9
JOURNAL J. Bacteriol. 185 (21), 6325-6330 (2003)
PUBMED 14563867
REFERENCE 2 (bases 1 to 360)
AUTHORS Julien,B.N.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Kosan Biosciences, 3832 Bay Center Place,
Hayward, CA 94545, USA

FEATURES

source
1..360
/organism="Myxococcus xanthus"
/mol_type="genomic DNA"
/strain="DZ1"
/db_xref="taxon:34"
complement(133..206)
/product="trna-gly"
165..206
misc_feature
/note="attB2; core bacterial attachment site for
bacteriophage Mx9"

ORIGIN

Query Match 100.0%; Score 42; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTGCTGCTTCTTACCAACTGAGTACCACGC 42

Db 165 GAGCTTGGGAAGCTCGTCTACCAACTGAGCTACACCGC 206

RESULT 2
AR619262
LOCUS AR619262 4575 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 657 from patent US 6833447.
ACCESSION AR619262
VERSION AR619262.1 GI:59723668
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4575)
AUTHORS Goldman,B.S., Hinkie,G.J., Slater,S.C. and Wiegand,R.C.
TITLE Myxococcus xanthus genome sequences and uses thereof
JOURNAL Patent: US 6833447-A 657 21-DEC-2004;
Monsanto Technology, LLC; St. Louis, MO
FEATURES
source 1..4575
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 42; DB 6; Length 4575;
Best Local Similarity 100.0%; Pred.No. 0.00013;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTCGTCTACCAACTGAGCTACACCGC 42
Db 1981 GAGCTTGGGAAGCTCGTCTACCAACTGAGCTACACCGC 2022

RESULT 3
LOCUS AY247757 4581 bp DNA linear PHG 27-OCT-2003
DEFINITION Bacteriophage Mx9 integration region, partial sequence.
ACCESSION AY247757
VERSION AY247757.2 GI:37992881
KEYWORDS Bacteriophage Mx9
SOURCE Bacteriophage Mx9
ORGANISM Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
P22-like viruses.

REFERENCE 1 (bases 1 to 4581)
AUTHORS Julien,B.
TITLE Characterization of the integrase gene and attachment site for the Myxococcus xanthus bacteriophage Mx9
JOURNAL J. Bacteriol. 185 (21), 6325-6330 (2003)
PUBMED 14563867
REFERENCE 2 (bases 1 to 4581)
AUTHORS Julien,B.N.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2003) Kosan Biosciences, 3832 Bay Center Place, Hayward, CA 94545, USA
REFERENCE 3 (bases 1 to 4581)
AUTHORS Julien,B.N.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2003) Kosan Biosciences, 3832 Bay Center Place, Hayward, CA 94545, USA
REMARK Sequence update by submitter
COMMENT On Oct 27, 2003 this sequence version replaced gi:32264340.
FEATURES
source Location/Qualifiers
1..4581
/organism="Bacteriophage Mx9"
/mol_type="genomic DNA"
/db_xref="taxon:235265"
/note="integration region"
607..2253
/gene="int"
607..2253
/gene="int"

gene

CDS

ORIGIN

/note="site specific recombinase"
/codon_start=1
/transl_table=11
/product="Int"
/protein_id="AAP78666.1"
/db_xref="GI:32264341"
/translation="MALGASDATNTNPSRLVQSVAAQPRATPWGYSASWYLLGRTATG
EYIVSSDAKKGHPNATAERLPTSPIDNALALEVARLVALQOQSATPPSSGRTFGA
VADDMLITEAKLVCPDNERRHLRMEALWGMTDELTVRVYKAHLGLLKEGGLPSA
ATVNVKSTGKRIIAQAINGEWPNVPGVLDREKEAKAERLTLTAACRAVLPHFR
ADRRREFLQVFLGPRGEEKALKKEDVDVEARTVIFRSNGRDTTKTGRERRVPVD
ELWPVLLDAQASPSDLVFPNAKGERQADTKMTVRLRTALSAAGVVVGWDYICRTQG
CGYRDVQSGGARQERRCPACDKRWASGRPKAVVWYGLRTAATLHRKAGCDPLVTKL
VLGHAAVTDDVYTHLDEDCYKALNKLSLAKPPPPHPOGSGDGGSPSGRNTYGE
GTWGGGLDLQHRARARARALPELPPRLAGGIPAPLLSVKQDVAAASLSVSTAKYQ
LLAAGVLPTVVWVQSRVKREDLDAYIAKATATGKRGKK"
2250..2735
/codon_start=1
/transl_table=11
/product="Orf2"
/protein_id="AAP78667.1"
/db_xref="GI:32264342"
/translation="MSRCPELEPRFSRLVARGEDACWLMOGRDLRNGYGRVAYNGR
DLSAHRVAYELATGTLPLPSVLVCHHCDNRCCNPAHLFAGTHADNMCAKGRQAR
GEGSTSKLTVAQWVELYRRRLAGESGPALAAAFGVSSRLVYLIAAGRKWAHTQPVAT
P"
2761..2970
/codon_start=1
/transl_table=11
/product="Orf7"
/protein_id="AAR07018.1"
/db_xref="GI:37993860"
/translation="MSRRPRNATRPEDSPDTLPDRFVLYVGRVPGWSVSPSIQAGP
FVSGRFGLLAAITKALKALARGR"
2967..3695
/codon_start=1
/transl_table=11
/product="Orf3"
/protein_id="AAP78668.2"
/db_xref="GI:37992882"
/translation="MSARERPILFSGAMVSAILAGKKTVTRVLKPKQPNPSPHDGGT
TWVFRADKGLHVPHGSGVGHLYVEERGLHCYPQGFQDLRLRLRETWALVRPAITDGPYV
EWADGALPKHEPKQDWLVLYRASVYDGAADTAERDFWRPSIHMPRWASRLTLEVV
SVRAERHALIDLPEEVLAEGLVREQDGDADDPPFRYGLPGHLTHRTGWAAAFANLW
ERINGEASNDANPWNWALSFRVP"
3692..3964
/codon_start=1
/transl_table=11
/product="Orf4"
/protein_id="AAP78669.1"
/db_xref="GI:32264344"
/translation="MSPLKFKAMGRGDAYEVRDGLLGQVLGVEMTKSVIGTNRAR
LSLVVRWNYRVEGIRITRGPYLTROKAAEVLASQYDAHQAKAVAP"
3961..4275
/codon_start=1
/transl_table=11
/product="Orf5"
/protein_id="AAP78670.1"
/db_xref="GI:32264345"
/translation="MTCAPAPRNNRGMPPSSSRVTPPQAEVLAFFVVTYDKHGYS
TADLCAHFQWKSNTNACQDMLTRLQDMHLITRAPKVIIRSLVVTESGREVAEWRREQE
QP"
4272..4581
/codon_start=1
/transl_table=11
/product="Orf6"
/protein_id="AAP78671.1"
/db_xref="GI:32264346"
/translation="MTCIGIHVTASGMPATKALRLDASEVORVANLRWAKGRGLADF
TGFVEAYLLKSARLFTHTPRLRGRQVWFGPAGCGSMALVYELLRPLRWERHLAELA
A"

CDS
/locus_tag="BT0737"
8906.._10573
/locus_tag="BT0737"
/codon_start=1
/transl_table=11
/product="formate--tetrahydrofolate ligase"
/protein_id="AA075844.1"
/db_xref="GI:29338042"
/translation="MKSDIEIARSVELKKIKQVAESIGIPREVENYGYIAKIPQL
IDEEKVKNLNLVLTATKAGIKGTIVSIGLALGNKIGKKAIVALREPSLGPFG
MKGAGAGGYAQLVPEKINLHFTGDFHAITSAHNMISALLDNYQNAQSGFLKEI
LWRRVLVDNDRSLRSTVGLSGNSITQESGDFITPASEIMAILCLSKDVEDLRRI
ENILGFTYDDQPTFKVLQAGAITVLKDAIHNLVOTTEGTAAFVHGPFANIAH
GNSILATKLANSFGDYVITTEAGFADLGAERFYNIKCRKSGLOPKLTVIVATAOGLK
MGGSVLDRIKPEBNEGLEKELRNLDKHINLRSGQTVVAFNFKFADITDEMEMLR
EHCEQVGATINNAFSGDGAVDARLVDTIENPSEPLRYTKEDSIOCKIEK
VATNLYGAVTYSIARNRIKLEKMGITHYPVCIAKTQYFSADPKTYGAVNPF
HKDIVNNGASIMIVAIAGEIMLMPGLPKEPQALHIDIIVDGEIGLS"
gene
/locus_tag="BT0738"
complement(10960..12240)
CDS
/locus_tag="BT0738"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="AA075845.1"
/db_xref="GI:29338043"
/translation="MKRDLIDFIEKEHORQLGIELIASENFSDQVMQMGSLT
NKYAEYQPKRYGCEVVDVQSEQIAIDRLKEIFGAENAVQPHSGAQANAFLAVL
NPGDFKPMGLNHLGSLVNTSGIITPCYNLANQETGRVDYDQMEEVALARKEP
MLIGGSAYSEWDYKRWREIADKAILMDMAHPAGLIAAGLLENPKVAHIVTS
TTHKTLRGRGVGVKMGDPFPMWCKTKPKGEIKMWSOLLDSAVPFGVQGPLEHVA
AKAVAFGEILQPEYKAYQVQNAALQAALIDRGFTIVSGTNDHSLMLDLRSKYP
DITGKVAEKALVSADITNKNMVPFDSRSFQTSGLRGLTPAITTRGAKEDLMIEIAE
MIETVLSVNEBEVIAQVARNVETMKKYPFLFAD"
gene
/locus_tag="BT0739"
complement(12461..13198)
CDS
/locus_tag="BT0739"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA075846.1"
/db_xref="GI:29338044"
/translation="WNILRNILFGWMAICTAFADQNRADKVERPNTKKINFGIKAG
FNSMFVSELKIKDVIDEVQNNYKIGYFGAIFMRFKHPIQPEASINYSKGEIT
FDKLGSQHPAIDPVASVQSVLHSDIFPILYGVNVYKPGPYGMGSIAPGPKLYLWGKQ
NEBITTFDQKGIHEKLYPLNYSVIGVGNISRIFFDPRYEQGIGNISVIYDNNIN
SDGSTGVSPILPRRDSALSPSGFPL"
gene
/locus_tag="BT0740"
complement(13215..13754)
CDS
/locus_tag="BT0740"
/codon_start=1
/transl_table=11
/product="flavoredoxin"
/protein_id="AA075847.1"
/db_xref="GI:29338045"
/translation="MIYPLPAILVSCGKDESYNIITVAWTGICTNPPMCIYSVRPE
RHSYDIIKKNMEFVNLTKDMAFPDWCQVSRNRYKFEEMKLTGRCITVVSAPLI
BESPLICECRKVEIIVSLGSHDMFIADVNVNRRADNLNPFCTGFELAEANPLVYVHGG
YIDLGEIKGKGSVEKKK"
gene
/locus_tag="BT0741"
complement(13816..14277)
CDS
/locus_tag="BT0741"
complement(13816..14277)

Query Match 86.2%; Score 36.2; DB 1; Length 300162;
Best Local Similarity 92.7%; Pred. No. 0.011;
Matches 38; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
|||||
DB 258334 AGCTTGGGAAGCTAATGCTCTACCAACTGAGCTACTACCGC 258374
|||||

RESULT 5
AE017174
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBLISHED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
repeat_region
gene
CDS
repeat_region
gene
CDS

AE017174
Porphyromonas gingivalis W83 section 3 of 8 of the complete genome.
AE017174 AE015924
AE017174.1 GI:34396675
Porphyromonas gingivalis W83
Porphyromonas gingivalis W83
Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
Porphyromonadaceae; Porphyromonas.
1 (bases 1 to 301933)
Nelson,K., Fleishmann,R., DeBoy,R., Paulsen,I., Fouts,D., Eisen,J.,
Daugherty,S., Dodson,R., Durkin,A., Gwinn,M., Haft,D., Kolonay,J.,
Nelson,W., White,O., Mason,T., Tallon,L., Gray,J., Granger,D.,
Tettelin,H., Dong,H., Galvin,J., Duncan,M., Dewhirst,F. and
Fraser,C.
Complete Genome Sequence of the Oral Pathogenic Bacterium
Porphyromonas gingivalis Strain W83
J. Bacteriol. 185 (18), 5591-5601 (2003)
12949112
2 (bases 1 to 301933)
Nelson,K., Fleishmann,R., DeBoy,R., Paulsen,I., Fouts,D., Eisen,J.,
Daugherty,S., Dodson,R., Durkin,A., Gwinn,M., Haft,D., Kolonay,J.,
Nelson,W., White,O., Mason,T., Tallon,L., Gray,J., Granger,D.,
Tettelin,H., Dong,H., Galvin,J., Duncan,M., Dewhirst,F. and
Fraser,C.
Direct Submission
Submitted (29-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
Location/Qualifiers
1..301933
/organism="Porphyromonas gingivalis W83"
/mol_type="genomic DNA"
/strain="W83"
/db_xref="taxon:242619"
86..1423
/insertion_seq="IS1.4"
complement(233..1318)
/locus_tag="PG0549"
complement(233..1318)
/locus_tag="PG0549"
/codon_start=1
/transl_table=11
/product="ISPgl, transposase"
/protein_id="AAQ65741.1"
/db_xref="GI:34396676"
/translation="MAYQSKNTDEHVTADALLSKRYKAONDFLNQVRLIDWRPIR
TLINKYTKRQNAICAPAYDVILLFKMLLETWYNSLDCALSERINDSITFSRFLGLK
MEVSPDSTISFRSALTGLMDKLAQFNKSRHHISVREGVLVDSLAVETPHK
PNGTITIEVADDREDNREAEKEEDYQKQVRRKRGTDPEARVYKQKRYHYGKK
HCLTNVQGIQKVITTAANSDTKBFIPLLOGANIQQGTAVLADKGYACGNRSYLQT
HHLQDMHKAQRNALTEEEKQNKAIPIRSITTIERTPGSIRRWPHGRCRYGLAK
THTQNLISIAFNLYRTPGIINSSSLG"
1557..2446
/note="miniature inverted-repeat transposable element"
/transposon="MITE_464.3"
complement(2613..5435)
/locus_tag="PG0553"
complement(2613..5435)
/locus_tag="PG0553"
/note="similar to GB:M63420, GB:J03261, SP:P15018,
PID:178415, PID:187141, PID:34362, and PID:575489;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="extracellular protease, putative"
/protein_id="AAQ65742.1"
/db_xref="GI:34396677"
/translation="MNKFKSLQLQGLAAVFSMATALTASAQISFGGPELFSFSSRSAG
THSFDDAMTIRLTPDNFNPEDLIAQRWSQDRGPRVRIQGVIPVDVDFASKASHISSI

		/locus_tag="PG0558"		
		/codon_start=1		
		/transl_table=11		
		/product="purine nucleoside phosphorylase I, inosine and guanosine-specific"		
		/protein_id="AAQ65747.1"		
		/db_xref="GI:34396682"		
		/translation="MKLQSQHVHEAASFSSRLPGDAKTAIIILGSLGELAEKIENT VTPYNEIPIHFAQATVAGKHGNTIGGILGCTPVVMQGRFHYEGYSMDQVTFPIRMVK LIGINELFVSNAAAGTNTISFKVQDLMIICDHINNPLNPLIGNMDFGVRFPDMTRAY DREFTKAGIAQELNI PVKEGVYGLTGPSTYETPAEYKFWGVQGDGDAIGMSTVPEVI VARTGIRVFGMSVLTNEGYPGADDFVNDQDVIRAANAASERKMGAI FARLIAAV" complement (10597..11742)		
gene		/locus_tag="PG0559"		
		/complement (10597..11742)		
CDS		/locus_tag="PG0559"		
		/codon_start=1		
		/transl_table=11		
		/product="chlorohydrolase family protein"		
		/protein_id="AAQ65748.1"		
		/db_xref="GI:34396683"		
		/translation="MAVYPLGQLQCHTSAMTIFRGYGGDLRLMDWLENWIWPVEAQM EDVYWGSKLACLEMKSGTTFADLWYHTLATARAEBMGLRAVLSSSTLDFRGDDBER ARIDRSCYALKEAFCSYSDRLOFQPSGVGHAIYTVSGEQLQFCHRFANEKNVLIHLHS ETEGEVDCIAKFGTTPVRYLKHGILSLQLILAHSLWDDMDLLAAGCKVKNP ASNMKLAGYRPHYDEMRRKRGTVIGLGTGDCSSNNLMIIAKKLAAFILGKAWRSAT AVKATDIYESATVDGARINGTDTVIAPGRADLCILRDIIPMTPCHPNIFISNLVYSA NSSAVDTTIVDGKILMRGRKVPGEBAIILAGAAEAAYKLMKAAHKTHE" complement (11719..11964)		
gene		/locus_tag="PG0560"		
		/complement (11719..11964)		
CDS		/locus_tag="PG0560"		
		/note="identified by Glimmer2; putative"		
		/codon_start=1		
		/transl_table=11		
		/product="hypothetical protein"		
		/protein_id="AAQ65749.1"		
		/db_xref="GI:34396684"		
		/translation="MQPGYNGHGRGIYFIGNMTANRRSLNHTVALDPDILDTDFQOC LPQYVAQCSILLKKEGATGFPFALYAPSTVSSVLYMR" complement (11946..13301)		
gene		/locus_tag="PG0561"		
		/complement (11946..13301)		
CDS		/locus_tag="PG0561"		
		/note="similar to GB:M96364, PID:1033197, PID:144480, and SP:Q59335; identified by sequence similarity; putative"		
		/codon_start=1		
		/transl_table=11		
		/product="peptidase, M20/M25/M40 family"		
		/protein_id="AAQ65750.1"		
		/db_xref="GI:34396685"		
		/translation="NNIRDIYIRENEARFLEDLFAIRIPSVSAKSHKPDQRCAEHW RDHLQVGNQKAEEVQTPGNVYVAERINDPKAKTILVYHYDVMPEPLELWKSSEF EPVIRDHIWARGADDKGGQMIQKVGFETALAGLVQCNVKFLPGEEBIEIGSTNLEA FCRAHKEMLADVIIVSDTFSWYSAETPSLTITGLRGLAYMEWTVIGNRDLHSGFGGA VAPNINELKLSJAVITDSTIRFIDNGRITIPHYDDVVPVLSDEERQIAQVFPSEKYYCOALDI DTVFCEAGYSTIERNSCRSPFQVCGIWWGTGEGATVLPKSAKAYKVSRLVANQDHE KTSQMFADYIRSAVPHKIHKVYTPHLGGEAYLCPIDLPAYKAAEEACTTAFGRKLPAV" complement (13057..14480)		
Query Match		86.2%; Score 36.2; DB 1; Length 301933;		
Best Local Similarity		92.7%; Pred. No. 0.011;		
Matches		38; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
QY	2	ACGTTGGGAAGCTCGTGCTCTACCACTGAGCTACCAACCGC	42	
DB	248859	ACGTTGGGAAGCTGATGCTCTACCACTGAGCTACTACCGC	248899	
RESULT 6				
AE001788/c				
LOCUS	AE001788	9890 bp	DNA	linear
DEFINITION	Thermotoga maritima MSB8 section 100 of 136 of the complete genome.			
ACCESSION	AE001788	AE000512		


```

/transl_table=11
/product="oxidase-related protein"
/protein_id="AAD36412.1"
/db_xref="GI:4981900"
/translations="MFAVLSSNEVRKRGVRLGRDLVFWRDSLGKVALEDFCVH
RRALRSQVINDRIOCPPHGPVDCNGVRLLI PALGKNVQVDPDRFVNSVPVYKNN
IWLMSGESEPEKEPFEDDIDEDPAXAFRELMVNPFFRAVENQLDVHLEFVHRTT
IIGNKTLVHGPVWIDESFIFYVNEVDQRQVRPEELSGEESVYLSEFIFPNL
WQNHISEGTRVAVAFFVDRQKTMILYRFGVMTGLKPDVDSIARLSMPLNRIILHOD
RRVVEQDIRKDLVQSDLPIMEFRKRLYKEKLLIDLFQGGQ"
7903..8376
/locus_tag="TM1341"
7903..8376
/locus_tag="TM1341"
/locus_tag="TM1341"
/notes="similar to percent identity: 0.00; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD36413.1"
/db_xref="GI:4981901"
/translations="MSGVTKYDPTSWRPKHLDLLLSAVIAFSLPAVAGWLYDLISGLVI
PMLIYVSLAWGLVKLRGVVGCTPPKPKPVYVNLVILTALPFAVOARPRREV
DRGVLLTALLWAPVNASFEIILMLYLFDSWDLYPEKRLRYELQACSSPFSWG"
8385..8594
/locus_tag="TM1342"
8385..8594
/locus_tag="TM1342"
/locus_tag="TM1342"
/notes="similar to percent identity: 0.00; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD36414.1"
/db_xref="GI:4981902"
/translations="MFWTKFLHVASPEGPFGVAFVLTMTISGFIHIVWRQSGNMVFT
FIPHFLNLGPGFWTGISVPLYLR"
8647..9036
/locus_tag="TM1343"
/notes="This region contains an authentic point mutation,
causing a premature stop, and is not the result of a
sequencing artifact; This region contains an authentic
frame shift and is not the result of a sequencing
artifact; similar to PID:1088462 PID:1848058"
complement(9194..9267)
/locus_tag="tRNA-Met-1"
complement(9194..9267)
/locus_tag="tRNA-Met-1"
/product="tRNA-Met"
/notes="codon recognized: AUG"
9373..9445
Query Match 82.4%; Score 34.6; DB 1; Length 9890;
Best Local Similarity 90.2%; Pred. No. 0.059;
Matches 37; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACACCGC 42
|||||
DB 9413 AGCTTGGGAAGCTGTGTCTTACCACTGAGCTACACCGC 9373

ECOGLYW 76 bp DNA linear BCT 26-APR-1993
E.coli glyW gene, a duplicate gene for gly-tRNA-3, major species.
J01624
J01624.1 GI:146227
glyW gene; transfer RNA; transfer RNA-Gly.
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 76)

```

```

AUTHORS Tucker,S.D., Gopalakrishnan,A.S., Bollinger,R., Dowhan,W. and
Murgola,E.J.
TITLE Molecular mapping of glyW, a duplicate gene for tRNA3Gly of
Escherichia coli
JOURNAL J. Bacteriol. 152 (2), 773-779 (1982)
PUBMED 6290450
COMMENT Original source text: Escherichia coli DNA.
GlyW was found to be within about 300 bp of pgsA [1]. The order of
genes in this region is uvrc, pgsA, glyW, flai. Comparison of the
order of determined restriction sites with the sites predicted from
the nucleotide sequence of gly-tRNA indicates that the direction of
transcription of glyW is counterclockwise on the circular E.coli
map.

FEATURES
source
1..76
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
<1..>76
/product="tRNA-Gly"
/notes="gly-tRNA-3 (glyW)"
ORIGIN 10 bp upstream from AluI site; about 41.9 min on K12 map.
Query Match 77.1%; Score 32.4; DB 1; Length 76;
Best Local Similarity 85.7%; Pred. No. 0.59;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACACCGC 42
|||||
DB 42 GACCTTGGCAAGTGTGTCTTACCAACTGAGCTATTCGCG 1

RESULT 8
ECOTRG3/c 76 bp tRNA linear BCT 15-APR-1994
LOCUS E.coli Gly-tRNA-3 with mutation Gly-tRNA-ins.
DEFINITION
ACCESSION K01549
VERSION K01549.1 GI:174426
KEYWORDS transfer RNA; transfer RNA-Gly.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 76)
Squires,C. and Carbon,J.
Normal and mutant glycine transfer RNAs
Nature New Biol. 233 (43), 274-277 (1971)
4941781
PUBMED 4941781
COMMENT Original source text: Escherichia coli tRNA.
Contributed on tape April 1983 by M.Sprinzl & D.H.Gauss; from their
entry 0730 in Nucleic Acids Res. 11, r1-r54 (1983). Ins =
insensitive to tryptone. The authors further suggest a model in
which three copies of a gene specifying Gly-tRNA-3 exist in close
proximity. These copies are suggested to be identical. [1] was
unable to detect 4-thiouridine in either the wild-type or
Gly-tRNA-ins, although a consistent low yield of oligonucleotides
containing 'us' suggests that this base may be modified.

FEATURES
source
1..76
/organism="Escherichia coli"
/mol_type="tRNA"
/db_xref="taxon:562"
1..76
/product="tRNA-Gly"
/notes="codon recognized: GGC; Gly-tRNA-3 (NAR: 0730)"
16 /mod_base=d
modified_base
17 /mod_base=d
modified_base
20 /mod_base=d
modified_base
34 /mod_base=d
variation
/notes="g in wild-type; t in Gly-tRNA-ins"

```



```

/db_xref="GI:49081054"
/translation="MRKIQARSQRPTTTPQSKLARASLHLFAAQRSSRRPERILP
QKEGRPPPPPNATERASSARDPGRYRETAGLERQNEASKQDQEQGAQRKK
GPKALSSSTQATWTFMGAGNETRDPDLGKIVLYQLSYRSRVTTCCPFGNRTFDAG
"
misc_difference 185
misc_difference 403
misc_difference 478
/replace="a"
/replace="g"
/replace="ta"

ORIGIN
Query Match 77.1%; Score 32.4; DB 11; Length 480;
Best Local Similarity 85.7%; Pred. No. 0.5;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTACCACTGAGCTACCACCGC 42
|||
Db 391 GACCTTGGCAAGATCGTGCTCTACCACTGAGCTATTCCCGC 432

RESULT 12
AF274339
LOCUS
DEFINITION Pseudomonas syringae pv. glycinea thermoresponsive genetic locus,
mutant PG4180.617 sequence.
ACCESSION AF274339
VERSION AF274339.1 GI:11037462
KEYWORDS
SOURCE Pseudomonas syringae pv. glycinea
ORGANISM Pseudomonas syringae pv. glycinea
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 490)
Ullrich,M.S., Schergaut,M., Boch,J. and Ullrich,B.
Temperature-responsive genetic loci in the plant pathogen
Pseudomonas syringae pv. glycinea
JOURNAL Microbiology 146 (Pt 10), 2457-2468 (2000)
PUBMED 11021922
AUTHORS Ullrich,M.S., Schergaut,M., Boch,J. and Ullrich,B.
DIRECT SUBMISSION
TITLE Submitted (02-JUN-2000) Ecophysiology, Max Planck Institute for
terrestrial Microbiology, Karl-von-Frisch-Strasse, Marburg 35043,
Germany
FEATURES
source 1..490
/organism="Pseudomonas syringae pv. glycinea"
/mol_type="genomic DNA"
/db_xref="taxon:318"
/notes="pathovar: glycinea"
misc_feature 1..490
/notes="thermoreponsive genetic locus, mutant PG4180.617;
similar to Erwinia chrysanthemi Orf1 possibly encoding a
cyclase involved in polyketide synthesis"

ORIGIN
Query Match 77.1%; Score 32.4; DB 1; Length 490;
Best Local Similarity 85.7%; Pred. No. 0.49;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTACCACTGAGCTACCACCGC 42
|||
Db 348 GACCTTGGCAAGTCTGTCTCTACCACTGAGCTATTCCCGC 389

RESULT 13
ECTRNGCL/c
LOCUS
DEFINITION E.coli glyW, cyst, and leuZ genes for tRNA-Gly, tRNA-Cys and
tRNA-Leu.
X52789

```

```

X52789.1 GI:43166
cyst gene; glyW gene; leuZ gene; transfer RNA gene; transfer
RNA-Cys; transfer RNA-Gly; transfer RNA-Leu.
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 500)
Komine,Y., Adachi,T., Inokuchi,H. and Ozeki,H.
Genomic organization and physical mapping of the transfer RNA genes
in Escherichia coli K12
J. Mol. Biol. 212 (4), 579-598 (1990)
PUBMED 2184240
FEATURES
source 1..500
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/map="42 min (2003Kb)"
/clone="340"
115..190
/gene="glyW"
115..190
/gene="glyW"
/product="tRNA-Gly"
245..318
/gene="cyst"
245..318
/gene="cyst"
/product="tRNA-Cys"
331..417
/gene="leuZ"
331..417
/gene="leuZ"
/product="tRNA-Leu"

ORIGIN
Query Match 77.1%; Score 32.4; DB 1; Length 500;
Best Local Similarity 85.7%; Pred. No. 0.49;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTCTACCACTGAGCTACCACCGC 42
|||
Db 156 GACCTTGGCAAGTCTGTCTCTACCACTGAGCTATTCCCGC 115

RESULT 14
HIHP1ATT
LOCUS
DEFINITION Haemophilus influenzae genes for tRNA-gly, tRNA-leu and tRNA-lys
and HPI host attachment site.
X53782
VERSION X53782.1 GI:43576
KEYWORDS bacteriophage attachment site; transfer RNA; transfer RNA-Gly;
transfer RNA-Leu; transfer RNA-Lys.
SOURCE Haemophilus influenzae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.
REFERENCE 1 (bases 1 to 636)
Hauser,M.A. and Socca,J.J.
AUTHORS Submitted (11-JUL-1990) Socca J.J., Dept. of Biochemistry, Johns
Hopkins Univ., School of Hygiene, 615 N. Wolfe Street, Baltimore,
MD 21205, USA
TITLE Location of the host attachment site for phage HPI within a cluster
of Haemophilus influenzae tRNA genes
JOURNAL Nucleic Acids Res. 18 (17), 5305 (1990)
PUBMED 1698282
REFERENCE 2 (bases 1 to 636)
Socca,J.J.
AUTHORS Direct Submission
TITLE Submitted (11-JUL-1990) Socca J.J., Dept. of Biochemistry, Johns
Hopkins Univ., School of Hygiene, 615 N. Wolfe Street, Baltimore,
MD 21205, USA
JOURNAL Data kindly reviewed (08-OCT-1990) by Socca J.
COMMENT Location/Qualifiers
FEATURES

```

```

source      1. .636
/organism="Haemophilus influenzae"
/mol_type="genomic DNA"
/strain="Rd001"
/db_xref="taxon:727"
/clone="pHPC102 and subclones"
complement(125. .143)
/notes="pot. transcriptional terminator"
misc_feature 161. .342
/notes="HPI attB/attP common core"
complement(161. .236)
trna         /product="trna-Lys"
complement(287. .373)
trna         /product="trna-Leu"
330. .347
misc_feature /notes="HPI attB sequence"
complement(378. .453)
trna         /product="trna-Gly"
486
misc_feature /notes="transcriptional start"
complement(494. .499)
promoter     /notes="pot. -10 region"
complement(516. .521)
promoter     /notes="pot. -35 region"

ORIGIN
Query Match      77.1%; Score 32.4; DB 1; Length 636;
Best Local Similarity 85.7%; Pred. No. 0.48;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGTACCACCGC 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 GACCTTGGCAAGTCGTCTTACCAACTGAGTATTCCCGC 453

RESULT 15
ECGLYWG      890 bp DNA linear BCT 07-JUL-2002
LOCUS        E. coli glyW region, a duplicated gene for trna-Gly-3.
DEFINITION   X03239 M4191
ACCESSION    X03239.1 GI:41604
VERSION      Insertion sequence; inverted repeat; transfer RNA-Gly; unidentified
KEYWORDS     reading frame.

SOURCE        Escherichia coli
ORGANISM      Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE    1 (bases 110 to 185)
AUTHORS      Tucker,S.D., Gopalakrishnan,A.S., Bollinger,R., Dowhan,W. and
              Murgola,E.J.
TITLE        Molecular mapping of glyW, a duplicate gene for trna3Gly of
              Escherichia coli
JOURNAL      J. Bacteriol. 152 (2), 773-779 (1982)
PUBMED      6290450
REFERENCE    2 (bases 1 to 890)
AUTHORS      Tucker,S.D. and Murgola,E.J.
TITLE        Sequence analysis of the glyW region in Escherichia coli
JOURNAL      Biochimie 67 (9), 1053-1057 (1985)
PUBMED      3002500
COMMENT      Data kindly reviewed (27-MAR-1986) by S. Tucker.
FEATURES
source      1. .890
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
<1. .10
repeat_region /insertion_seq="ISI"
11. .14
promoter      /notes="pot. -35 sequence"
31. .37
promoter      /notes="pot. -10 sequence"
53. .57
repeat_unit   /notes="imp. inverted repeat A"

```

```

repeat_unit 63. .68
/notes="imp. inverted repeat A"
110. .185
trna         /product="trna-Gly"
/notes="mature transfer RNA-Gly-3"
215. .232
misc_signal  /note="pot. transcription termination loop structure"
215. .220
repeat_unit  /note="imp. inverted repeat B"
227. .232
repeat_unit  /note="imp. inverted repeat B"
456. .463
repeat_unit  /note="inverted repeat C"
468. .475
repeat_unit  /note="inverted repeat C"
491. .496
promoter     /note="pot. -35 sequence"
514. .520
promoter     /note="pot. -10 sequence"
608. .838
CDS          /note="unnamed protein product; unidentified reading
              frame"
/codon_start=1
/transl_table=11
/protein_id="CAA26984.1"
/db_xref="GI:41605"
/db_xref="UniProt/Swiss-Prot:P06979"
/translation="MKTGPLNSELEWLDLITKYNTDHAILDVAELDGLLTAVLSSP
              QSEIPEQMLVAVWGADYVFRWASEKEMTRAL"

ORIGIN
Query Match      77.1%; Score 32.4; DB 1; Length 890;
Best Local Similarity 85.7%; Pred. No. 0.47;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGTACCACCGC 42
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 GACCTTGGCAAGTCGTCTTACCAACTGAGTATTCCCGC 110

```

Search completed: January 10, 2006, 18:22:11
Job time : 2188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: January 10, 2006, 10:49:30 ; Search time 484 Seconds
(without alignments)
578.341 Million cell updates/sec

Title: US-10-645-818-5

Perfect score: 42

Sequence: 1 gacttggaagctctgtct.....accaactgagctaccaccgc 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002as.*

7: Geneseq2002bs.*

8: Geneseq2003as.*

9: Geneseq2003bs.*

10: Geneseq2003cs.*

11: Geneseq2003ds.*

12: Geneseq2004as.*

13: Geneseq2004bs.*

14: Geneseq2005as.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	42	12	ADM36323 Bacteriop
2	42	100.0	74	12	ADM36321 Myxococcu
3	42	100.0	257	12	ADM36309 Bacteriop
4	42	100.0	360	12	ADM36306 Bacteriop
5	42	100.0	360	12	ADM36307 Bacteriop
6	42	100.0	1647	12	ADM36304 Bacteriop
7	42	100.0	4575	14	ACL64194 M. xanthu
8	32.4	77.1	258	12	ADL02418 DNA encod
9	32.4	77.1	549	11	ABD16247 Pseudomon
10	32.4	77.1	657	4	AAS51204 Salmonell
11	32.4	77.1	690	11	ABD16506 Pseudomon
12	32.4	77.1	1960	13	ADT05165 Haemophil
13	32.4	77.1	5436	13	ADT05456 Haemophil
14	32.4	77.1	5559	2	AAQ55260 Restricti
15	32.4	77.1	5559	2	AAT96101 Pseudomon
16	32.4	77.1	5559	2	AAT99212 5.6 kb E1
17	32.4	77.1	5559	2	AAT90945 E11 genom
18	32.4	77.1	5559	2	AAT06486 5.6 kb Ec
19	32.4	77.1	10763	2	AAV39837 Pseudomon

20	32.4	77.1	10763	2	AAx99368
21	32.4	77.1	28871	13	ADT05539
22	32.4	77.1	31147	4	AAF28538
23	32.4	77.1	63563	4	AAF28546
24	32.4	77.1	110000	2	AAT42063_00
25	32.4	77.1	110000	2	AAT42063_01
26	32.4	77.1	110000	2	AAT42063_17
27	32.4	77.1	110000	10	ACF67367_38
28	32.4	77.1	110000	10	ACF65388_09
29	32.4	77.1	308766	13	ADT05738
30	32.4	77.1	349980	13	ADT05648
31	29.8	71.0	339	10	ADP03634
32	29.8	71.0	30312	10	ACF65384_5
33	29.8	71.0	40681	6	ABA92787_6
34	29.8	71.0	110000	10	ACF67367_11
35	28.2	67.1	282	9	ADA29439
36	28.2	67.1	4927	3	AAA81466
37	28.2	67.1	7035	14	ACL64415
38	28.2	67.1	12531	4	AAS59530
39	28.2	67.1	12531	8	ACF64459
40	28.2	67.1	110000	3	AAA81489_2
41	28.2	67.1	110000	14	AE835724_4
42	28.2	67.1	110000	14	AE839174_07
43	28.2	67.1	110000	14	AE839175_25
44	28.2	67.1	110000	14	AE842401_25
45	28.2	67.1	110000	14	AE842737_11

ALIGNMENTS

RESULT 1

ADM36323

ID ADM36323 standard; DNA; 42 BP.

XX

AC ADM36323;

XX

DT 03-JUN-2004 (first entry)

XX

DE Bacteriophage Mx9 int gene attP core site SEQ ID NO:5.

XX

KW bacterial; bacteriophage Mx9; integrase; enzyme; int gene;

KW Mx9 transformation system; Mx9 integrase; molecular biology;

KW drug development; attP; attachment site; gene; ds.

XX

OS Bacteriophage Mx9.

XX

PN W02004018635-A2.

XX

PD 04-MAR-2004.

XX

PF 20-AUG-2003; 2003WO-US026413.

XX

PR 21-AUG-2002; 2002US-0405196P.

XX

PA (KOSA-) KOSAN BIOSCIENCES INC.

 PI | Julien B; |

CC attachment site recognised by the integrase protein, which may be the
 CC same as the first attachment site. Also described is a bacterial host
 CC cell produced by the method described above. The methods and compositions
 CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC exemplification of the present invention.

CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC Myxococcus xanthus trnagly nucleotide sequence, which is used in the
 CC exemplification of the present invention.

CC the bacteriophage Mx9 int gene attP core attachment site, which is used
 CC in the exemplification of the present invention.

SQ Sequence 42 BP; 9 A; 14 C; 11 G; 8 T; 0 U; 0 Other;
 Query Match 100.0%; Score 42; DB 12; Length 42;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 42
 |||||
 Db 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 42
 |||||

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 42
 |||||
 Db 42 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 1

RESULT 2

ADM36321/c
 ID ADM36321 standard; RNA; 74 BP.

AC ADM36321;

DT 03-JUN-2004 (first entry)

DE Myxococcus xanthus trnagly nucleotide sequence.

KW bacterial; Chrysoperla carnea; integrase; enzyme; int gene;
 KW Mx9 transformation system; Mx9 integrase; molecular biology;
 KW drug development; trnagly; ss.

OS Myxococcus xanthus.

PH Key Location/Qualifiers

FT stem_loop 10..25

FT /*tag= a

FT stem_loop 27..43

FT /*tag= b

FT stem_loop 49..65

FT /*tag= c

PN WO2004018635-A2.

PD 04-MAR-2004.

PF 20-AUG-2003; 2003WO-US026413.

PR 21-AUG-2002; 2002US-0405196P.

PS (KOSA-) KOSAN BIOSCIENCES INC.

PA Julien B;

PI WPI; 2004-226826/21.

PT Modifying a DNA of a bacterial cell comprises using an integrase gene and
 PT an attachment site for Myxococcus xanthus bacteriophage Mx9, useful in
 PT the fields of molecular biology and drug development.

PS Disclosure; Fig 4; 37pp; English.

XX The present invention describes a method for modifying a DNA of a
 CC bacterial cell having in its genome a first attachment site recognised by
 CC a protein with Mx9 integrase activity comprising introducing an Mx9
 CC transformation system into the cell, the system comprising a gene
 CC encoding a protein with Mx9 integrase activity protein operably linked to
 CC a promoter active in the host cell, and a DNA vector comprising a second
 CC attachment site recognised by the integrase protein, which may be the
 CC same as the first attachment site. Also described is a bacterial host
 CC cell produced by the method described above. The methods and compositions

CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC Myxococcus xanthus trnagly nucleotide sequence, which is used in the
 CC exemplification of the present invention.

SQ Sequence 74 BP; 11 A; 22 C; 25 G; 0 T; 16 U; 0 Other;

Query Match 100.0%; Score 42; DB 12; Length 74;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 42
 |||||
 Db 42 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGTACACCGC 1

RESULT 3

ADM36309

ID ADM36309 standard; DNA; 257 BP.

AC ADM36309;

DT 03-JUN-2004 (first entry)

DE Bacteriophage Mx9 integrase attB2 site nucleotide SEQ ID NO:7.

KW bacterial; bacteriophage Mx9; integrase; enzyme; int gene;
 KW Mx9 transformation system; Mx9 integrase; molecular biology;
 KW drug development; attB2; gene; ds.

OS Bacteriophage Mx9.

PN WO2004018635-A2.

PD 04-MAR-2004.

PF 20-AUG-2003; 2003WO-US026413.

PR 21-AUG-2002; 2002US-0405196P.

PS (KOSA-) KOSAN BIOSCIENCES INC.

PI Julien B;

PP WPI; 2004-226826/21.

PT Modifying a DNA of a bacterial cell comprises using an integrase gene and
 PT an attachment site for Myxococcus xanthus bacteriophage Mx9, useful in
 PT the fields of molecular biology and drug development.

PS Disclosure; SEQ ID NO 7; 37pp; English.

XX The present invention describes a method for modifying a DNA of a
 CC bacterial cell having in its genome a first attachment site recognised by
 CC a protein with Mx9 integrase activity comprising introducing an Mx9
 CC transformation system into the cell, the system comprising a gene
 CC encoding a protein with Mx9 integrase activity protein operably linked to
 CC a promoter active in the host cell, and a DNA vector comprising a second
 CC attachment site recognised by the integrase protein, which may be the
 CC same as the first attachment site. Also described is a bacterial host
 CC cell produced by the method described above. The methods and compositions
 CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC bacteriophage Mx9 int gene attB2 site nucleotide sequence, which is used
 CC in the exemplification of the present invention.

SQ Sequence 257 BP; 54 A; 81 C; 82 G; 40 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 12; Length 257;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 42
 |||||
 Db 115 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 156

RESULT 4
 ADM36306
 ID ADM36306 standard; DNA; 360 BP.
 AC ADM36306;
 XX
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Bacteriophage Mx9 integrase attB1 site nucleotide SEQ ID NO:3.
 XX
 KW bacterial; bacteriophage Mx9; integrase; enzyme; int gene;
 KW Mx9 transformation system; Mx9 integrase; molecular biology;
 KW drug development; attB1; gene; ds.
 XX
 OS Bacteriophage Mx9.
 XX
 PN WO2004018635-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2003; 2003WO-US026413.
 XX
 PR 21-AUG-2002; 2002US-0405196P.
 XX
 PA (KOSA-) KOSAN BIOSCIENCES INC.
 XX
 PI Julien B;
 XX
 DR WPI; 2004-226826/21.
 XX
 XX Modifying a DNA of a bacterial cell comprises using an integrase gene and
 PT an attachment site for Myxococcus xanthus bacteriophage Mx9, useful in
 PT the fields of molecular biology and drug development.
 XX
 PS Disclosure; SEQ ID NO 3; 37pp; English.
 XX
 CC The present invention describes a method for modifying a DNA of a
 CC bacterial cell having in its genome a first attachment site recognised by
 CC a protein with Mx9 integrase activity comprising introducing an Mx9
 CC transformation system into the cell, the system comprising a gene
 CC encoding a protein with Mx9 integrase activity protein operably linked to
 CC a promoter active in the host cell, and a DNA vector comprising a second
 CC attachment site recognised by the integrase protein, which may be the
 CC same as the first attachment site. Also described is a bacterial host
 CC cell produced by the method described above. The methods and compositions
 CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC bacteriophage Mx9 int gene attB1 site nucleotide sequence, which is used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 360 BP; 65 A; 104 C; 126 G; 65 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 42; DB 12; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 42
 |||||
 Db 163 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 204

RESULT 5
 ADM36307
 ID ADM36307 standard; DNA; 360 BP.
 XX
 AC ADM36307;

XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Bacteriophage Mx9 integrase attB2 site nucleotide SEQ ID NO:4.
 XX
 KW bacterial; bacteriophage Mx9; integrase; enzyme; int gene;
 KW Mx9 transformation system; Mx9 integrase; molecular biology;
 KW drug development; attB2; gene; ds.
 XX
 OS Bacteriophage Mx9.
 XX
 PN WO2004018635-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2003; 2003WO-US026413.
 XX
 PR 21-AUG-2002; 2002US-0405196P.
 XX
 PA (KOSA-) KOSAN BIOSCIENCES INC.
 XX
 PI Julien B;
 XX
 DR WPI; 2004-226826/21.
 XX
 XX Modifying a DNA of a bacterial cell comprises using an integrase gene and
 PT an attachment site for Myxococcus xanthus bacteriophage Mx9, useful in
 PT the fields of molecular biology and drug development.
 XX
 PS Disclosure; SEQ ID NO 4; 37pp; English.
 XX
 CC The present invention describes a method for modifying a DNA of a
 CC bacterial cell having in its genome a first attachment site recognised by
 CC a protein with Mx9 integrase activity comprising introducing an Mx9
 CC transformation system into the cell, the system comprising a gene
 CC encoding a protein with Mx9 integrase activity protein operably linked to
 CC a promoter active in the host cell, and a DNA vector comprising a second
 CC attachment site recognised by the integrase protein, which may be the
 CC same as the first attachment site. Also described is a bacterial host
 CC cell produced by the method described above. The methods and compositions
 CC of the present invention are useful in the fields of molecular biology
 CC and drug development, and particularly in transforming host bacterial
 CC cells using a bacteriophage Mx9 system. The present sequence represents a
 CC bacteriophage Mx9 int gene attB2 site nucleotide sequence, which is used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 360 BP; 68 A; 110 C; 124 G; 58 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 42; DB 12; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3.7e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 42
 |||||
 Db 165 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACCACCGC 206

RESULT 6
 ADM36304
 ID ADM36304 standard; DNA; 1647 BP.
 XX
 AC ADM36304;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Bacteriophage Mx9 int gene SEQ ID NO:1.
 XX
 KW bacterial; bacteriophage Mx9; integrase; enzyme; int gene;
 KW Mx9 transformation system; Mx9 integrase; molecular biology;
 KW drug development; gene; ds.
 XX
 OS Bacteriophage Mx9.
 XX

PH Key Location/Qualifiers
FT CDS 1..1647
FT /*tag= a
FT /product= "Mx9 integrase"

PN WO2004018635-A2.

XX 04-MAR-2004.

XX 20-AUG-2003; 2003WO-US026413.

XX 21-AUG-2002; 2002US-0405196P.

XX (KOSA-) KOSAN BIOSCIENCES INC.

XX Julien B;

XX WPI; 2004-226826/21.

XX P-PSDB; ADM36305.

XX Modifying a DNA of a bacterial cell comprises using an integrase gene and
PT an attachment site for Myxococcus xanthus bacteriophage Mx9, useful in
PT the fields of molecular biology and drug development.

XX Claim 5; SEQ ID NO 1; 37pp; English.

XX The present invention describes a method for modifying a DNA of a
CC bacterial cell having in its genome a first attachment site recognised by
CC a protein with Mx9 integrase activity comprises introducing an Mx9
CC transformation system into the cell, the system comprising a gene
CC encoding a protein with Mx9 integrase activity protein operably linked to
CC a promoter active in the host cell, and a DNA vector comprising a second
CC attachment site recognised by the integrase protein, which may be the
CC same as the first attachment site. Also described is a bacterial host
CC cell produced by the method described above. The methods and compositions
CC of the present invention are useful in the fields of molecular biology
CC and drug development, and particularly in transforming host bacterial
CC cells using a bacteriophage Mx9 system. The present sequence encodes
CC bacteriophage Mx9 integrase, which is used in the exemplification of the
CC present invention.

XX Sequence 1647 BP; 292 A; 528 C; 572 G; 255 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 12; Length 1647;
Best Local Similarity 100.0%; Pred. No. 5.2e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACCGC 42

Db 1394 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACCGC 1435

RESULT 7

ACL64194

ID ACL64194 standard; DNA; 4575 BP.

XX ACL64194;

XX 02-JUN-2005 (first entry)

XX M. xanthus DNA fragment, seq id 657.

XX Transgenic plant; DNA replication; gene regulation; gene expression; ds.

XX Myxococcus xanthus.

XX US6833447-B1.

XX 21-DEC-2004.

XX 10-JUL-2001; 2001US-00902540.

XX 10-JUL-2000; 2000US-0217883P.

XX (MONS) MONSANTO TECHNOLOGY LLC.
PA Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX WPI; 2005-028716/03.

XX New substantially purified Myxococcus xanthus nucleic acid molecule

PT encoding a nitrite reductase, useful for determining gene expression,

PT identifying mutations in a gene of interest, and for constructing

XX mutations in a gene of interest.

XX Example 1; SEQ ID NO 657; 25pp; English.

XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 1-1849 represent a
CC set of 1849 contig and singleton sequences comprising coding sequences,
CC DNA replication elements, promoters and other regulatory elements from
CC the genome of the bacterium Myxococcus xanthus. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO

XX Sequence 4575 BP; 636 A; 1611 C; 1559 G; 769 T; 0 U; 0 Other;

Query Match 100.0%; Score 42; DB 14; Length 4575;

Best Local Similarity 100.0%; Pred. No. 6.5e-07;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACCGC 42

Db 1981 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACCAACCGC 2022

RESULT 8

ADL02418

ID ADL02418 standard; DNA; 258 BP.

XX ADL02418;

XX 06-MAY-2004 (first entry)

XX DNA encoding a M. catarrhalis protein #104.

XX ds; gene; Moraxella catarrhalis; infection.

XX Moraxella catarrhalis.

XX US6673910-B1.

XX 06-JAN-2004.

XX 04-APR-2000; 2000US-00540236.

XX 08-APR-1999; 99US-0128416P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL;

XX WPI; 2004-178127/17.

XX P-PSDB; ADL04338.

XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for

PT preparing a composition for diagnosing, preventing or treating infection

PT caused by Moraxella catarrhalis.

XX Disclosure; SEQ ID NO 104; 429pp; English.

CC The invention relates to an isolated nucleic acid encoding an Moraxella
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
 CC composition for diagnosing, preventing or treating infection caused by
 CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
 CC catarrhalis protein.
 XX
 SQ Sequence 258 BP; 82 A; 57 C; 44 G; 75 T; 0 U; 0 Other;

Query Match 77.1%; Score 32.4; DB 12; Length 258;
 Best Local Similarity 85.7%; Pred. No. 0.003;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAGCTCGTCTTACCACTGAGCTACCAACCGC 42
 DB 113 GACCTTGGCAAGGTCGTCTTACCACTGAGCTATTCCCGC 154

RESULT 9
 ABD16247
 ID ABD16247 standard; DNA; 549 BP.
 XX AC ABD16247;
 DT 29-JUL-2004 (first entry)
 XX Pseudomonas aeruginosa polynucleotide #14851.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.
 XX Pseudomonas aeruginosa.
 OS US6551795-B1.
 PN 22-APR-2003.
 PD 18-FEB-1999; 99US-00252991.
 PF 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 DR WPI; 2003-615309/58.
 DR P-PSDB; AB082676.
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX Disclosure; SEQ ID NO 14851; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABD197-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 549 BP; 122 A; 166 C; 166 G; 95 T; 0 U; 0 Other;

Query Match 77.1%; Score 32.4; DB 11; Length 549;
 Best Local Similarity 85.7%; Pred. No. 0.0036;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAGCTCGTCTTACCACTGAGCTACCAACCGC 42
 DB 478 GACCTTGGCAAGGTCGTCTTACCACTGAGCTATTCCCGC 519

RESULT 10
 AAS51204/c
 ID AAS51204 standard; DNA; 657 BP.
 XX AC AAS51204;
 DT 13-FEB-2002 (first entry)
 XX Salmonella typhimurium cellular proliferation inhibitory sequence #102.

XX Antisense; ss; prokaryotic cellular proliferation; antibiotic;
 KW antibacterial; drug design.
 XX Salmonella typhimurium.
 OS WO200170955-A2.
 PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.
 PF 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.

(ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyakind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

PS Claim 1; SEQ ID NO 3781; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence is an antisense oligonucleotide of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 657 BP; 167 A; 140 C; 164 G; 183 T; 0 U; 3 Other;

Query Match 77.1%; Score 32.4; DB 4; Length 657;
 Best Local Similarity 85.7%; Pred. No. 0.0037;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACACCGC 42
|||
Db 579 GACCTTGGCAAGGTCGTGTCTTACCAACTGAGCTATTCCCGC 538

RESULT 11
ABD16506
ID ABD16506 standard; DNA; 690 BP.
XX
AC ABD16506;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #15110.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR P-PSDB; ABO82935.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 15110; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 690 BP; 149 A; 217 C; 214 G; 110 T; 0 U; 0 Other;

Query Match 77.1%; Score 32.4; DB 11; Length 690;
Best Local Similarity 85.7%; Pred. No. 0.0038;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACACCGC 42
|||
Db 44 GACCTTGGCAAGGTCGTGTCTTACCAACTGAGCTATTCCCGC 85

ID ADT05165 standard; DNA; 1960 BP.
XX
AC ADT05165;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 201.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
XX contig.
XX
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
DR WPI; 2004-662422/64.
XX
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
XX Example 1; SEQ ID NO 201; 88pp; English.
XX
XX The invention comprises nucleotide sequences (genes) from the genome of a
XX nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
XX sequences of the invention are useful for treating or preventing NTHi
XX bacterial infections of the middle ear and/or nasopharynx. The present
XX nucleic acid represents an NTHi contig sequence of the invention.
XX
XX Sequence 1960 BP; 603 A; 371 C; 394 G; 592 T; 0 U; 0 Other;

Query Match 77.1%; Score 32.4; DB 13; Length 1960;
Best Local Similarity 85.7%; Pred. No. 0.0048;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTCGTGTCTTACCAACTGAGCTACACCGC 42
|||
Db 760 GACCTTGGCAAGGTCGTGTCTTACCAACTGAGCTATTCCCGC 801

RESULT 13
ADT05456/C
ID ADT05456 standard; DNA; 5436 BP.
XX
AC ADT05456;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 492.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
XX contig.
XX
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX

PA (CHIL-) CHILDRENS HOSPITAL INC.
XX Bakaletz LO, Munson RS, Dyer DW;
XX WPI; 2004-662422/64.
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX Example 1; SEQ ID NO 492; 88pp; English.
XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi contig sequence of the invention.
XX Sequence 5436 BP; 1705 A; 1150 C; 884 G; 1691 T; 0 U; 6 Other;
SQ Query Match 77.1%; Score 32.4; DB 13; Length 5436;
Best Local Similarity 85.7%; Pred. No. 0.0061;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
DB 5153 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 5112
RESULT 14
AAQ55260
ID AAQ55260 standard; DNA; 5559 BP.
XX AC AAQ55260;
XX 25-MAR-2003 (revised)
DT 28-JUL-1994 (first entry)
XX XX
XX Restriction fragment comprising gene activating sequence.
XX Gene activation; gafa; transcriptional activator; activator; sensor;
KW regulation; induction; biocontrol agents; crop protection; ss.
XX Pseudomonas fluorescens.
XX Key Location/Qualifiers
FT CDS 210..1688
FT /*tag= a
FT /label= ORF 1.
FT CDS 1906..3633
FT /*tag= b
FT /label= ORF 2.
FT CDS 4616..4691
FT /*tag= c
FT /label= glyW
FT /note= "tRNA coding sequence."
FT CDS 4731..5318
FT /*tag= d
FT /label= ORF 3.
XX WO9401561-A1.
XX 20-JAN-1994.
XX 02-JUL-1993; 93WO-US006300.
XX 02-JUL-1992; 92US-00908284.
XX (CIBA) CIBA GEIGY AG.
XX Gaffney TD, Lam ST, Hill DS, Stein JT, Howell CR, Becker J;
PI Ligon JM;
XX

DR WPI; 1994-035067/04.
XX Gene activating sequences, partic. from pseudomonas - used for inducing
PT expression of genes which are latent or expressed at low levels in target
PT organisms.
XX Disclosure; Page 71-74; 87pp; English.
XX The gene activating sequence is contained within an 11 kilobase EcoRI
CC fragment (E11) from P. fluorescens. DNA sequence analysis of E11 has
CC identified five open reading frames (ORF), of which ORF5 and ORF 2 share
CC significant homology with numerous activator and sensor components
CC respectively of bacterial two-component regulatory systems. The gene
CC activating sequences can be used in microorganisms to induce the
CC production of gene products and secondary metabolites. They are used
CC particularly for production of bacterial biocontrol agents for use in
CC inhibiting plant pathogens. This sequence is a fragment of E11. See also
CC AAQ55261. (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 5559 BP; 1194 A; 1743 C; 1637 G; 984 T; 0 U; 1 Other;
SQ Query Match 77.1%; Score 32.4; DB 2; Length 5559;
Best Local Similarity 85.7%; Pred. No. 0.0061;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
DB 4650 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 4691
RESULT 15
AAT96101
ID AAT96101 standard; DNA; 5559 BP.
XX AC AAT96101;
XX 31-MAR-1998 (first entry)
XX XX
XX Pseudomonas fluorescens lemA gene fragment.
XX lemA gene; gafa kinase; biocontrol factor; production; chitinase;
KW gelatinase; pyrrolnitrin; cyanide; Pseudomonas; biological control;
KW pathogenic fungus; ss.
XX Pseudomonas fluorescens.
XX US5686282-A.
XX 11-NOV-1997.
XX 02-JUN-1995; 95US-00459701.
XX 20-AUG-1990; 90US-00570184.
XX 02-JUL-1992; 92US-00908284.
XX 01-JUL-1993; 93US-00087636.
XX 08-JUN-1994; 94US-00258261.
XX 08-AUG-1994; 94US-00287442.
XX (NOVS) NOVARTIS FINANCE CORP.
XX Lam ST, Gaffney TD;
XX WPI; 1997-558141/51.
XX Pseudomonas fluorescens lemA gene - useful for activating gene expression
PT in Pseudomonas.
XX Example 5; Col 41-45; 35pp; English.
XX The present sequence is a Pseudomonas fluorescens lemA gene fragment. The
CC P. fluorescens lemA gene, or lemA derivatives with enhanced gafa kinase
CC activity, can be used to increase the production of biocontrol factors,
CC e.g. chitinase, gelatinase, pyrrolnitrin and cyanide in Pseudomonas,

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2006, 16:47:24 ; Search time 3793 Seconds
(without alignments)
518.075 Million cell updates/sec

Title: us-10-645-818-5
Perfect score: 42
Sequence: 1 gagcttggaagctcgtcgt.....acaaactgagctaccaccgc 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	32.4	77.1	760	10	CL688860 PRI014c.P
C 2	32.4	77.1	773	10	CL692030 PRI01581
C 3	32.4	77.1	782	9	BZ564906 pac82-164
C 4	32.4	77.1	806	9	BZ551245 pac81-60
C 5	32.4	77.1	811	10	CL664382 PRI0146d
C 6	32.4	77.1	1250	9	BZ563705 pac82-164
C 7	32.4	77.1	1393	9	BZ577553 mah2 5462
C 8	32.4	77.1	1635	9	BZ575755 mah2_4612
C 9	30.8	73.3	556	7	CN694592 E0352609-
C 10	30.8	73.3	665	5	BU702692 UI-M-FCO-
C 11	30.8	73.3	680	6	CF725482 UI-M-GZO-
C 12	30.8	73.3	719	6	CA320878 UI-M-FWO-
C 13	30.8	73.3	729	6	CF732823 UI-M-CZO-
C 14	29.8	71.0	427	11	AL390545 Leishmani
C 15	29.8	71.0	503	2	BE430009 TAS006.B0
C 16	28.8	68.6	922	9	BH163135 ENT1B32TF
C 17	28.2	67.1	696	9	BH371106 AG-ND-175
C 18	27.6	65.7	692	10	AG391506 Mus muscu
C 19	27.2	64.8	606	8	CX096269 EHAGP57TR
C 20	27.2	64.8	708	9	CC822117 MSBP12B8
C 21	26.8	63.8	180	9	AZ608686 IM0433L04
C 22	26.8	63.8	312	9	AZ945284 2M0206B12

23	26.8	63.8	946	9	BH137767
C 24	26.6	63.3	440	9	BH373173
C 25	26.6	63.3	740	9	BH380524
C 26	26.6	63.3	743	10	CW874974 ehs2kd10-
C 27	26.6	63.3	748	10	CW880257 ehs2kd10-
C 28	26.6	63.3	942	9	BH138539 ENTOR81TF
C 29	26.4	62.9	593	9	AZ293626 RPCI-23-9
C 30	26.4	62.9	623	9	AZ103897 RPCI-23-2
C 31	26.2	62.4	191	9	AZ578476 23H03 Sho
C 32	26	61.9	540	10	CW016252 ZMMBLCO00
C 33	26	61.9	545	9	CC386674 PUHMO28TD
C 34	26	61.9	554	9	AQ988749 RPCI-23-3
C 35	26	61.9	627	9	BH899419 Ote00486
C 36	26	61.9	657	9	CE118795 tigr-988-
C 37	26	61.9	659	10	CW860959 sHM2hd7-1
C 38	26	61.9	720	10	BX164483 dario rer
C 39	26	61.9	744	10	AG372472 Mus muscu
C 40	26	61.9	865	1	AV988406 AV988406
C 41	26	61.9	1297	10	AG429724 Mus muscu
C 42	25.8	61.4	547	10	CW873528 she2hd7-8
C 43	25.6	61.0	375	5	BY022217 BY022217
C 44	25.6	61.0	489	9	AZ291650 RPCI-23-1
C 45	25.6	61.0	574	9	AZ632374 IM0486P24

ALIGNMENTS

RESULT 1
CL688860/c
LOCUS
DEFINITION
CL688860 760 bp DNA linear GSS 09-JUL-2004
PRI014c.P06.2 - PRI014c.BR (760) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION
CL688860 GI:50197938
VERSION
GSS.
KEYWORDS
Pristionchus pacificus
SOURCE
Pristionchus pacificus
ORGANISM
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE
1 (bases 1 to 760)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppABD: an AcedB database for the nematode satellite organism
Pristionchus pacificus
JOURNAL
Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED
14681447
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES

Location/Qualifiers
1..760
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
clone lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 77.1%; Score 32.4; DB 10; Length 760;
Best Local Similarity 85.7%; Pred. No. 0.088; Indels 0; Gaps 0;
Matches 36; Conservative 0; Mismatches 6;

QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 259 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 218
 |||||

RESULT 2
 CL692030/c
 LOCUS
 DEFINITION
 CL692030.1 GI:50213938
 CL692030
 GSS.
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 773)
 Srinivasan J., Otto G.W., Kahlow U., Geisler R. and Sommer R.J.
 AppadB: an AcedB database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 14681447
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.

FEATURES
 source
 1..773
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
 Query Match 77.1%; Score 32.4; DB 10; Length 773;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 331 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 290
 |||||

RESULT 3
 BZ564906/c
 LOCUS
 DEFINITION
 BZ564906.1 GI:27191664
 BZ564906
 GSS.
 Pseudomonas aeruginosa
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 782)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..782
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 806;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 331 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 290
 |||||

RESULT 4
 BZ551245/c
 LOCUS
 DEFINITION
 BZ551245.1 GI:27154826
 BZ551245
 GSS.
 Pseudomonas aeruginosa
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 806)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 20622216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 806;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 331 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 290
 |||||

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

COMMENT
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 20622216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.

FEATURES
 source
 1..782
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="2-164"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 2-164 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 782;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 378 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 337
 |||||

RESULT 5
 BZ551245/c
 LOCUS
 DEFINITION
 BZ551245.1 GI:27154826
 BZ551245
 GSS.
 Pseudomonas aeruginosa
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 806)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 20622216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 806;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 378 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 337
 |||||

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 806;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 378 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 337
 |||||

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

ORIGIN
 Query Match 77.1%; Score 32.4; DB 9; Length 806;
 Best Local Similarity 85.7%; Pred. No. 0.088;
 Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GAGCTTGGGAAGCTCGTCTCTACCAACTGAGCTACACCGC 42
 |||||
 Db 378 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCCGC 337
 |||||

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..806
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone_lib="pacsl-60 3056"
 /note="clinical isolate 1-60 Whole genomic shotgun
 library."

Db 384 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 343

RESULT 5
CL664382/c
LOCUS
DEFINITION
CL664382 811 bp DNA linear GSS 09-JUL-2004
PRI0146d.G08 - PRI0146d.B21 (811) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CL664382.1 GI:50154798
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 811)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism

JOURNAL
PUBLISHED
COMMENT

Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

Location/Qualifiers

1..811
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 77.1%; Score 32.4; DB 10; Length 811;
Best Local Similarity 85.7%; Pred. No. 0.088;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 42

DB 706 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 665

RESULT 6
BZ563705/c
LOCUS
DEFINITION
BZ563705 1250 bp DNA linear GSS 17-DEC-2002
pac2-164_4385.y2 pac2-164 Pseudomonas aeruginosa genomic clone
pac2-164_4385, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BZ563705.1 GI:27188137
GSS.
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 1250)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1..1250
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac2-164_4385"
/clone_lib="pac2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN

Query Match 77.1%; Score 32.4; DB 9; Length 1250;
Best Local Similarity 85.7%; Pred. No. 0.093;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 42

DB 201 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 160

RESULT 7

BZ577553

LOCUS

DEFINITION

BZ577553

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1393)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1..1393

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="MSH"

/db_xref="taxon:287"

/clone="meh2 5462"

/clone_lib="msh"

/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN

Query Match 77.1%; Score 32.4; DB 9; Length 1393;
Best Local Similarity 85.7%; Pred. No. 0.094;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAGCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 42

DB 431 GACCTTGGCAAGTCTGCTCTACCAACTGAGCTATTCCGCG 472

RESULT 8
BZ575755
LOCUS
DEFINITION
msh2_4612.x1 msh Pseudomonas aeruginosa genomic clone msh2_4612,
genomic survey sequence.
ACCESSION
BZ575755
VERSION
BZ575755.1 GI:27210816
KEYWORDS
GSS:
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1635)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1..1635
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="M5H"
/db_xref="taxon:287"
/clone_lib="msh2_4612"
/notes="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 77.1%; Score 32.4; DB 9; Length 1635;
Best Local Similarity 85.7%; Pred. No. 0.095;
Matches 36; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
|||||
DB 444 GACCTTGGCAAGCTCGTCTTACCAACTGAGCTATTCGCG 485
|||||

RESULT 9
CN694592
LOCUS
DEFINITION
E0352G09-5 NIA Mouse E10.5 whole embryo cDNA library (long) Mus
musculus cDNA clone NIA:E0352G09 IMAGE:30863984 5', mRNA sequence.
ACCESSION
CN694592
VERSION
CN694592.1 GI:47463341
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 556)
AUTHORS
Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Basse,y,U.C.,
Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L.,
Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A.,
Vescovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A.,
D'Urso,M., Kelsoe,J., Hide,W. and Ko,M.S.
TITLE
Transcriptome analysis of mouse stem cells and early embryos
PLOS Biol. 1 (3), 410-419 (2003)
COMMENT
Contact: Dawood B. Dudekula

Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: E0352 row: G column: 09
Seq primer: M13 Reverse
High quality sequence stop: 556
POLYA=No.

FEATURES
source
Location/Qualifiers
1..556
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:E0352G09-5"
/db_xref="taxon:10090"
/clone="NIA:E0352G09 IMAGE:30863984"
/tissue_type="whole embryo including extraembryonic
tissues at 10.5-days postcoitum"
/dev_stage="E10.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E10.5 whole embryo cDNA library
(long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 8 embryos at 10.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGCGCGCCCTTTT-3'] from
2ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Loner-linker Ld-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.4Kb. The library was
constructed by Yulan Piao."

ORIGIN
Query Match 73.3%; Score 30.8; DB 7; Length 556;
Best Local Similarity 83.3%; Pred. No. 0.36;
Matches 35; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 GAGCTTGGGAAGCTCGTCTTACCAACTGAGCTACACCGC 42
|||||
DB 438 GTGCTTGGGAAGCAAGCGCTCTACCAACTGAGCTACCCCTC 479
|||||

RESULT 10
BUT02692
LOCUS
DEFINITION
UI-M-FCO-bz2b-p-24-0-UI.r1 NIH BMAP_FCO Mus musculus cDNA clone
IMAGE:6401543 5', mRNA sequence.
ACCESSION
BUT02692
VERSION
BUT02692.1 GI:23627751
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 665)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

The following repetitive elements were found in this cDNA sequence: 7-103, >PB1D10#SINE/Alu (matched complement) 45-185, >B4#SINE/B4 (matched complement)

FEATURES

source

FEATURES

source

Location/Qualifiers

1. .665
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6401543"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FC0"

/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 73.3%; Score 30.8; DB 5; Length 665;
Best Local Similarity 83.3%; Pred. No. 0.37;
Matches 35; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTGCTCTACCACTGAGCTACACCGC 42
DB 139 GTGCTTGGGAAGCAGCGCTCTACCACTGAGCTACACCGC 180

RESULT 11

CF725482

LOCUS UI-M-GZO-cj-m-c-18-0-UI.r1 NIH BMAP_G20 Mus musculus cDNA clone
DEFINITION IMAGE:30603233 5', mRNA sequence.

ACCESSION CF725482.1 GI:37599650

VERSION CF725482

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 680)

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

The following repetitive elements were found in this cDNA sequence: 1-116, >URR1B#DNA/MER1 type 300-442, >B1 MM#SINE/Alu (matched complement) 346-442, >PB1D10#SINE/Alu (matched complement) 384-524, >B4#SINE/B4 (matched complement)

Seq primer: pYX-5.

Location/Qualifiers

1. .680
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30603233"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP G20"

/notes="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTATTGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 73.3%; Score 30.8; DB 6; Length 680;
Best Local Similarity 83.3%; Pred. No. 0.37;
Matches 35; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAGCTTGGGAAGCTGCTCTACCACTGAGCTACACCGC 42
DB 478 GTGCTTGGGAAGCAGCGCTCTACCACTGAGCTACACCGC 519

RESULT 12

CA320878

LOCUS

DEFINITION

UI-M-FW0-ccb-f-24-0-UI.r1 NIH BMAP_FW0 Mus musculus cDNA clone

IMAGE:6817273 5', mRNA sequence.

CA320878

ACCESSION CA320878.1 GI:24538976

VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 719)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

Search completed: January 10, 2006, 19:25:38
Job time : 3798 secs

This Page Blank (uspto)